AAMRL-TR-86-014

DEVELOPMENT OF SIMULATED DIRECTIONAL AUDIO FOR COCKPIT APPLICATIONS (U)

THEODORE J. DOLL
JEFFREY M. GERTH
WILLIAM R. ENGELMAN
DENNIS J. FOLDS

SYSTEMS ENGINEERING LABORATORY
GEORGIA TECH RESEARCH INSTITUTE
GEORGIA INSTITUTE OF TECHNOLOGY

JANUARY 1986

Approved for Public Release; Distribution is Unlimited

HARRY G. ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY
AEROSPACE MEDICAL DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433-6573

NOTICES

TECHNICAL REVIEW AND APPROVAL

AAMRL-TR-86-014

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Director, Human Engineering Division
Armstrong Aerospace Medical Research Laboratory

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION
Unclassified

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE

3. DISTRIBUTION/AVAILABILITY OF REPORT
Approved for public release; distribution is unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AAMRL-TR-86-014

6a. NAME OF PERFORMING ORGANIZATION
Georgia Institute of Technology
Georgia Tech Research Institute

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION
AAMRL/HEA

6c. ADDRESS *(City, State and ZIP Code)*
Atlanta, GA 30332

7b. ADDRESS *(City, State and ZIP Code)*
Wright-Patterson AFB OH 45433-6573

8a. NAME OF FUNDING/SPONSORING ORGANIZATION

8b. OFFICE SYMBOL *(If applicable)*

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
Subcontract 1054-9A in support of prime contract F33615-82-C-0513

8c. ADDRESS *(City, State and ZIP Code)*

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61101F	7184	26	D2

11. TITLE *(Include Security Classification)*
SEE REVERSE

12. PERSONAL AUTHOR(S)
Doll, Theodore J., Gerth, Jeffrey M., Engelman, William R. and Folds, Dennis J.

13a. TYPE OF REPORT
Technical

13b. TIME COVERED
FROM Nov 84 TO Jan 86

14. DATE OF REPORT *(Yr., Mo., Day)*
1986 January

15. PAGE COUNT
230

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Auditory Localization; Directional Cues; Visually-Coupled Systems; Head-Coupled Control/Display Systems; Three-Dimensional Sound

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

The long-term objective of this work is to develop techniques for conveying accurate spatial information via audio signals delivered to the listener through headphones or earphones. This project included three major activities: (1) an extensive review and synthesis of the research literature on auditory localization, (2) the design, fabrication, and evaluation of an apparatus for demonstrating simulated auditory localization (SAL), and (3) experimental research to determine characteristics of the audio signal, in the time and frequency domains, which enhance localization performance with simulated cues.

Previous research is reviewed which describes the cues involved in the perception of sound-source direction, both horizontally and vertically, when the head is stationary. Also reviewed is research on auditory distance perception, the roles of head movement and vision in auditory localization, the perception of auditory motion and volume, and the effects of noise on auditory localization. A feedback control model (continued on back)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☒ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
Gloria L. Calhoun

22b. TELEPHONE NUMBER *(Include Area Code)*
(513) 255-7595

22c. OFFICE SYMBOL
AAMRL/HEA

DD FORM 1473, 83 APR   EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

11. TITLE: Development of Simulated Directional Audio for Cockpit Applications (U)

20. ABSTRACT - continued

is presented, which integrates evidence derived from four different theoretical positions concerning the effects of head movement and vision on auditory localization. Possible applications of SAL technology in aircraft cockpits are outlined, and the potential benefits of such applications are discussed. Topic areas in auditory localization are identified where further basic research is needed to support the development of SAL into a valuable pilot aid.

The design, evaluation, and refinement of the SAL facility is described. An experimental test of the psychological fidelity of the SAL facility is summarized. The results show that the facility produces a high-fidelity simulation of normal, unaided auditory localization.

Two experiments are described in which listeners localized sounds on the basis of simulated cues delivered dichotically. Localization accuracy and response time were compared for: (1) nine different filtered noise stimuli, designed to make available some localization cues while eliminating others, and (2) 16 different intermittent noise stimuli, formed by manipulating burst duration, duty cycle, rise time, and the frequency composition of the noise. The stimulus characteristics had relatively little effect on localization speed and accuracy. It is concluded that, when the listener's head is free to move, localization performance with simulated cues is relatively insensitive to the character of the audio signal.

A long-term research and development plan for an electronic simulator of auditory localization cues is presented. The simulator would be suitable for airborne applications, have the capability to impress directional qualities on incoming signals in real time, and be able to compensate for head movements so as to make directional signals appear to be stationary in space. Alternative approaches are discussed and major technical issues are outlined.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

PREFACE

The research described in this report was performed by the Systems Engineering Laboratory of the Georgia Tech Research Institute (GTRI) of the Georgia Institute of Technology. The project was sponsored by the U.S. Air Force Armstrong Aerospace Medical Research Laboratory (AAMRL) under subcontract number 1054-9A with MacAulay-Brown, Inc., in support of prime contract number F33615-82-C-0513.

The principal investigator and project director for this was Dr. Theodore J. Doll of GTRI. The Air Force project technical monitor was Ms. Gloria Calhoun of the AAMRL Human Engineering Division, Visual Display Systems Branch (AAMRL/HEA).

The authors wish to thank the following individuals for their help in designing and building parts of the Simulated Auditory Localization (SAL) research facility: Professor W. Marshall Leach of the Georgia Tech School of Electrical Engineering and Mr. Harold F. Engler and Mr. Richard V. Morrison of GTRI. Thanks are due to Dr. Thomas A. Furness and Dr. Robert G. Eggleston of AAMRL/HEA and Dr. Bryce O. Hartman of the Air Force School of Aerospace Medicine for their support of this program. The authors also express their appreciation to Mr. Paul H. Cleveland and Mr. J. Craig Wyvill of GTRI for loaning the project valuable pieces of test equipment; and to Knowles Electronics Co. for providing miniature microphones free of charge.

TABLE OF CONTENTS

TABLE OF CONTENTS
(continued)

TABLE OF CONTENTS
(concluded)

LIST OF FIGURES

LIST OF FIGURES
(continued)

LIST OF FIGURES
(concluded)

LIST OF TABLES

LIST OF TABLES
(continued)

LIST OF TABLES
(concluded)

1.0 INTRODUCTION

1.1 Background

In military systems, particularly aircraft, the preponderance of information is presented to the human operator in the visual modality. During critical periods of high workload, this produces an overload in the visual channel which can reduce pilot and system effectiveness, safety, and survivability.

The visual overload problem could be alleviated by making better use of the auditory modality, particularly for information which humans customarily process aurally (c.f., Doll, Folds, & Leiker, 1984; Simpson, 1982; Werkowitz, 1981). For example, in the natural environment, sounds are often a cue as to the direction of important objects or sources of information. Upon hearing a novel sound, a human generally rotates his or her head and eyes toward the source and acquires it visually. Auditory information is thus a natural cue for visual acquisition of objects in the environment.

This research is directed toward developing ways to convey accurate spatial information via audio signals delivered to the listener through earphones or headphones. Some of the ways in which such simulated directional audio cues could be used to enhance the performance of pilots of military aircraft are the following:

1. Redirection of visual attention
   Particularly in Head-Coupled Control/Display systems, a natural, non workload-intensive method is needed to cue the pilot where to look.
2. Enhancement of situational awareness
   Directional audio cues could be used to help keep the pilot aware of aircraft attitude in low visibility conditions, and to warn of threats or closing terrain when visual attention is directed elsewhere.
3. Enhancement of communications and audio warnings
   Giving audio messages and radio communications each a different apparent direction could enhance their intelligibility and lessen pilot response time, especially in noise and jamming conditions.

1.2 Objectives

The specific objectives of this work were as follows:

1. Design, fabricate, and evaluate an apparatus for demonstrating simulated auditory localization (SAL), i.e. auditory localization based on simulated cues provided via headphones.
2. Perform a review and synthesis of the research literature which provides a basis for: (a) evaluating the feasibility and potential value of using SAL in the cockpit, and (b) for identifying areas where further research is needed to develop SAL into a valuable pilot aid.
3. Conduct experimental research to determine characteristics of the auditory stimulus, in the time and frequency domains, which enhance localization performance with simulated audio cues.

1.3 Summary and Conclusions

Section 2.0 of this report provides an extensive review of the research literature on auditory localization. The first part of the review focuses on the cues involved in the perception of sound-source direction and distance relative to the listener when the listener's head is stationary.

The experimental evidence indicates that at least five, and possibly six, cues play a role in the perception of the direction of a sound source relative to the listener. The cues are known to the extent that the general physical mechanism which produces each cue (e.g., the pinnae or outer ear) and the region of the spectrum in which each cue operates can be specified. The changes in the acoustic waveform at the eardrums or ear canal entrance as a function of source direction have been precisely measured for many listeners, and some common features have been identified. However, such waveforms, called head-related transfer functions (HRTF's) are complex. Consequently, there exists, to date, only a limited understanding of what features of the HRTF's are actually used by the ear and brain as cues to source direction.

In the study of auditory distance perception, it is useful to distinguish between absolute and relative cues. Absolute cues allow the listener to judge source distance in the absence of any additional information about the source, e.g., its spectral composition and direction. Absolute cues are, of necessity, binaural or involve head movement. Initial research suggests that the interaural time (or phase) difference (ITD) is the major absolute distance cue.

Relative cues allow the listener to judge the absolute distance of a sound source when the source characteristics (intensity and spectral composition) are known or assumed. For example, one can judge the distance of human speakers assuming that they are conversing at normal conversational intensity. Relative cues also allow a listener to judge whether a single, known source is closer or further away on two successive occasions. Further research is needed to identify all the cues for auditory distance perception, to determine how they interact, and to quantify how they vary as a function of source distance.

Section 2.5 reviews research on the roles of head movement and vision in auditory localization. The evidence deriving from each of four different theoretical perspectives is reviewed and integrated into a single model which encompasses all the findings. Auditory localization in the presence of head movement and vision is viewed as a multiple-level feedback control system. Visual information, prior knowledge, and feedback from motor activities (especially eye movements) provide important inputs which help to control auditory perception.

Sections 2.6 and 2.7 review the research on the perception of auditory motion and volume, and the effects of noise on auditory localization. Research in these areas is still in its infancy. Since these topics are crucial to applications of SAL in the cockpit, further research is urgently needed. Section 6.0, the long-term research and development plan, presents a program of proposed research in this topic area.

Sections 2.8 and 2.9 include a discussion of possible applications of SAL technology in aircraft cockpits, the potential benefits of such applications, and topic areas in auditory localization where further basic research is needed.

Another major objective of this project was to design, build, and evaluate a facility for producing SAL cues. Section 3.0 describes the SAL research facility, now in operation at the Georgia Tech Research Institute (GTRI). Section 4.0 reviews a series of tests and refinements which were made to the SAL facility to enhance and evaluate its psychological fidelity. It was concluded that the SAL facility produces a high fidelity simulation of normal, free-field auditory localization.

An experiment was conducted in order to compare localization performance in the SAL facility (i.e., with simulated cues) to that in normal, unaided localization. There was initially a small difference in localization accuracy (about 3 degrees RMS) which decreased to virtually zero after about one hour of practice. The average time to localize a sound source with simulated cues was slightly longer than with normal cues (about 3.4 sec versus 2.5 sec), and this difference was still present after an hour of practice.

Section 5.0 summarizes research which examined how the characteristics of auditory signals in the time and frequency domains affects their usefulness as directional signals in the cockpit. Two experiments were conducted in which human listeners localized sounds on the basis of simulated cues delivered via headphones. The first experiment compared localization performance in the SAL facility for high- and low-pass filtered noise signals. The cut-off frequencies of the stimuli were designed to pass portions of the spectrum associated with some localization cues, while rejecting portions associated with others. The second experiment used intermittent signals and examined the effects on localization performance of burst duration, duty cycle (repetition rate), and rise time for high- and low-frequency noise bursts. In contrast to most previous studies in which the listener's head remained stationary, stimulus characteristics had relatively little effect on localization speed and accuracy. It is concluded that modulation of the received sound produced by head movement allows the listener to judge the source direction, and that such modulation renders cues associated with the spectral and time-domain composition of the sound much less important for localization than when the head is stationary.

Although the overall effects of stimulus characteristics were small, the rise time and repetition rate of intermittent stimuli produced trends which have significance for the design of directional auditory displays. Specifically, shorter rise times (e.g., 1 msec) and higher repetition rates (2 Hz and greater) produce better localization performance than longer rise times and lower repetition rates.

Overall, the results of the experiments suggest that, when the listener can move the head, localization performance with simulated cues is relatively insensitive to the character of the audio signal. Thus localization performance should not be seriously degraded for nonoptimal audio signals such as tones and speech, as long as the listener's head is free to move.

The last section of this report (6.0) presents a long-term research and development plan for an electronic simulator which could convey highly accurate directional information by way of audio signals to a listener wearing earphones. The simulator will be suitable for airborne applications and will have the capability to impress directional qualities on incoming signals in real time. The simulator will also be able to alter the sound quality in real time in a manner coordinated with the listener's head movement. This is necessary in order to make a sound appear to be stationary in space as the listener moves the head, and/or to simulate a moving sound source. Section 6.2 examines alternative approaches for building a real-time directional synthesizer for audio signals, and Section 6.3 outlines the major technical issues to be resolved. Section 6.4 outlines a research and development plan for building and testing a prototype of the real-time directional synthesizer.

2.0 LITERATURE REVIEW AND RECOMMENDATIONS

2.1 Introduction

This section includes a review of the basic research literature on auditory localization and of one major attempt to use auditory localization in a human-machine system (Garner, 1949). The purpose of this review is three-fold. First, the review provides a basis for evaluating the feasibility of using simulated auditory localization (SAL) to provide directional cues in head-coupled control/display systems. A second purpose of this review is to identify other potential uses of SAL in human-machine systems, and to assess the feasibility of each such use. The third objective is to identify requirements for further basic research which would facilitate the development of SAL applications in human-machine systems.

The first topic addressed in this review is the acoustic cues which enable humans to perceive the location of a sound source when the listener's head is immobile. Specifically, the first three subsections of this review address cues to sound-source positions in the horizontal plane, the median plane, and cues to distance. The next three subsections address the effects of head and eye movement on auditory localization, the perception of auditory motion and volume, and the effects of noise on auditory localization. The final two subsections discuss possible applications of SAL in the cockpit and present recommendations for needed basic research on auditory localization. An annotated bibliography is presented in Appendix A.

Throughout this report, "horizontal plane" means the horizontal plane passing through the listener's interaural axis. The term "median" plane refers to the median sagittal plane, or the plane of symmetry of the body. The term "cue" refers to the physical characteristics of the proximal acoustic stimulus at the listener's ear canal entrance or eardrum which allow him or her to locate the source in space.

2.2 Localization in the Horizontal Plane

For many years, "duplex" theory dominated thinking and research on auditory localization. Duplex theory holds that two types of acoustic cues, interaural time difference (ITD) and interaural amplitude difference (IAD), account for localization in the horizontal plane. Even some relatively modern sources discuss ITD and IAD cues extensively, but make little or no mention of other important cues. The persistence of duplex theory is probably related to the fact that most auditory localization research has been performed with the stimulus presented via headphones. This method allows the experimenter to freely manipulate the proximal stimulus at each ear, but also eliminates cues due to head movement and reflections and resonances of the pinnae and torso.

It is only relatively recently that the cues necessary for localization in the horizontal plane have been understood well enough to obtain accurate localization performance from a listener wearing earphones. Garner (1949) summarized several unsuccessful attempts to simulate auditory localization for submarine applications as follows:

> "When the problem of auditory signaling is mentioned, usually the first thing that comes to mind is the use of some indication of lateral displacement of a sound source. If time, intensity, and phase differences between the two ears are the cues we use in localizing a sound source, then it should be simple to produce an apparent displacement of a sound source by stimulating the two ears differently in one of these respects. Unfortunately, it is not as simple as that. (p.212)
>
> ... Whatever the outcome of this type of research, it is clear that more research is needed before it will be possible to simulate good localization of sounds." (p. 213)

When ITD and IAD cues are simulated and presented dichotically, via earphones, the resulting sound is heard as if it were inside the head. Under these conditions, the sound can be made to subjectively appear to the listener as though it were located at various positions between his ears. However, the sound always remains subjectively within the head; hence this type of experiment is called lateralization. The ability to judge the corresponding position outside the head, i.e., to localize, is quite poor in these conditions (Mills, 1972).

Batteau and his colleagues (Batteau, 1967, Batteau, Plante, Spencer & Lyle, 1963; 1965) were apparently the first to demonstrate auditory localization performance using earphones with accuracy comparable to that achieved with the unaided ears. Previous researchers, dating back to the 19th Century, had speculated that the external ears, or pinnae, play a role in localization by altering the quality of the sound depending on its direction of origin (Shaw, 1982a). Bloch (1893) and McLean (1959) demonstrated that distortion of the pinnae impairs the ability to localize sounds (cited in Shaw, 1982b and Batteau, 1967, respectively). Angell and Fite (1901a,b) demonstrated that monaural localization is possible, although not with the accuracy of binaural stimulation, and they suggested that the pinna plays a role in localization (cited by Mills, 1972 and Gatehouse, 1982a).

The method used by Batteau et al. (1965) to demonstrate simulated auditory localization is shown in Figure 1. A listener was stimulated dichotically using semi-insertion-electrostatic headphones with the signals recorded from a pair of artificial pinnae. The pinnae were mounted on high-fidelity microphones, which were attached to a bar and separated so as to correspond to the width of a human head. These artificial "ears" were located in a separate room, acoustically isolated from the listener. The listener reported the apparent azimuth of a maraca shaken at various positions around the artificial ears. The listener's head was restrained during the testing. When the artificial pinnae were attached to the microphones, the listener was able to localize the maraca with excellent accuracy in both azimuth and elevation. Without the artificial pinnae, the listener's judgements were erratic and inaccurate. Interestingly, with the artificial pinnae attached to the microphones, the listener reported that the sound appeared subjectively to be located some distance outside the head, rather than in the head, as sounds presented via earphones are normally perceived.

Since the Batteau et al. (1965) demonstration, investigators have continued to analyze the cues necessary for auditory localization. More recent findings suggest that auditory localization is quite complex, depending on as many as six different physical cues whose influence varies depending on the location of the sound source relative to the listener's head and the frequency of the stimulus. A number of investigators have measured the transfer functions which describe how the spectrum of the free-field sound is related to that of the proximal stimulus at the ear canal entrance or at the eardrum (c.f., Flannery & Butler, 1981; Gardner, 1973; Mehrgardt & Mellert, 1977; Rodgers, 1981; Shaw, 1974b; Shaw & Teranishi, 1968; Weinrich, 1982). Shaw (1982a) compiled a family of curves representing the average transformation of sound pressure level (SPL) from the free field to the human eardrum for 100 subjects studied in 12 separate studies. These curves, called Head-Related Transfer Functions (HRTF's), are reproduced in Figure 2 and show

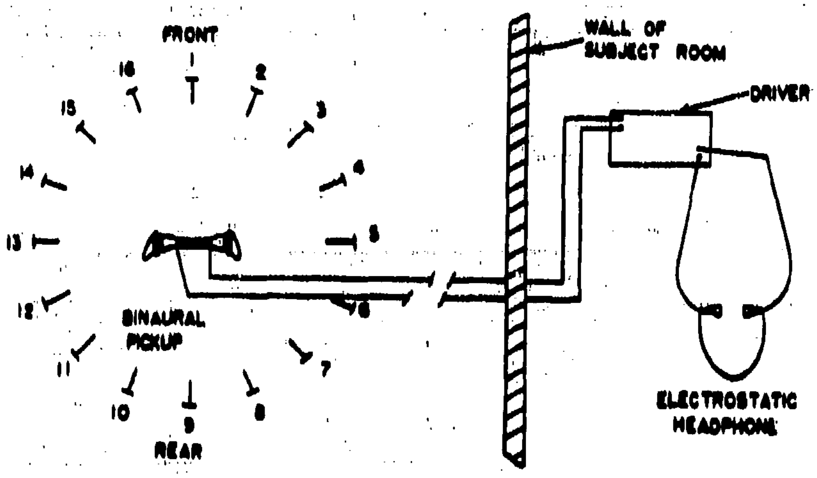


Figure 1. Method used by Batteau et al. (1965) to simulate auditory localization.

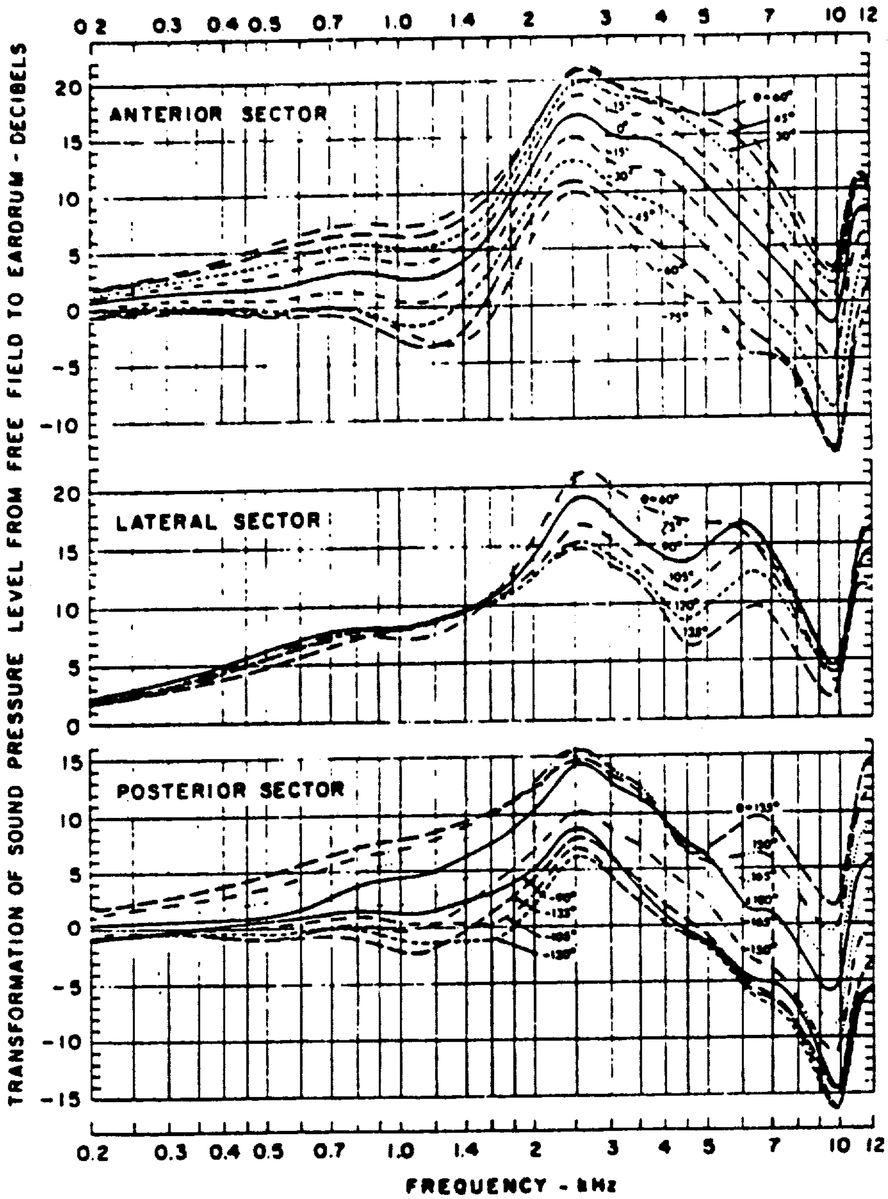


Figure 2. Average transformation of sound pressure level from free field to human eardrum as a function of frequency at 24 angles of incidence in the horizontal plane. (From Shaw, 1982b.)

SPL at the eardrum as a function of frequency for 24 angles of incidence. There are clear differences as a function of angle. Notice the substantial difference for sounds in the anterior and posterior directions at comparable angles. For example, the difference in SPL at 4 kHz for sounds located at 45 degrees and 135 degrees azimuth is about 10 db. Shaw concluded that the external ear produces an abundance of direction-dependent cues which could serve as the sole basis of auditory localization. However, it is well known that binaural localization is more accurate than monaural localization, hence interaural differences must also play an important role.

Searle, Braida, Davis & Colburn (1976) developed a model of the decoding process in auditory localization based on statistical decision theory, using data from studies of horizontal and vertical localization reported over a ten-year period. They conclude that there are six cues which the ear and brain uses to decode the direction of a sound, and that monaural pinna cues are of less importance than interaural pinna cues (see Section 2.3.1 for further discussion of this point). The six cues and their properties are listed in Table 1. Figure 3 defines the angular coordinates used in Table 1. Searle et al. (1976) used a nonlinear least-squares procedure to estimate the standard deviation of localization possible on the basis on each cue alone. The estimates, based on date from 18 experiments, are shown in Table 2. These estimates are for binaural localization in the horizontal plane with a broad-band white noise stimulus and a speaker array spanning 90°. Although ITD and IAD are among the most important cues, it is clear that pinna and torso cues also play a role.

Measurements of auditory acuity provide another kind of evidence that localization performance may be based on several different types of cues. Mills (1958) measured the minimum audible angle (MAA) for pure tones as a function of frequency and the direction of the initial sound. The listeners judged whether two successive sounds came from the same or different positions. A loudspeaker was sounded at an initial azimuth value, and then moved through a small angle to a second position. The subject's head was restrained. The MAA is that separation between the first and second speaker positions at which the subject correctly judged whether the second sound came from the left or right of the first sound on 75% of the trials. The results are shown in Figure 4. The MAA is about 1° for low frequency tones directly

TABLE 1

Auditory Localization Cues and Their Properties[a]

Cue number	Description	Assumed spatial dependence[b]	Useful frequency range	*A priori* spectral information required?
1	Interaural time delay	θ	20 Hz–12 kHz	No
2	Interaural amplitude difference arising from head shadow	θ	1–12 kHz	No
3	Monaural head shadow	θ	1–12 kHz	Yes
4	Interaural pinna amplitude response	θ, γ	4–12 kHz	No
5	Monaural pinna amplitude response	θ, γ	4–12 kHz	Yes
6	Amplitude response due to shoulder reflections	θ, γ	2–3 kHz	Yes

[a]Adapted from Searle et al. (1976).

[b]See Figure 3

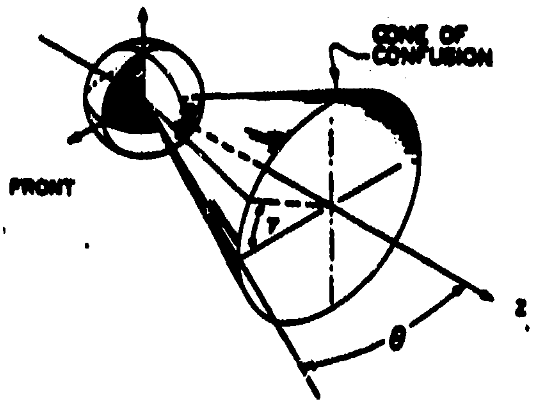


Figure 3. Spherical head model, showing the cone of confusion and defining the spherical coordinate system. (From Searle et al., 1976.)

TABLE 2

Estima[illegible] Standard Deviation of Localization (In Degrees) Attributable to Six Cues[1]

Cue Description		Standard Deviation of Localization[2]
Interaural time delay	3	5.6°
Interaural head shadow		
Monaural head shadow		
Interaural pinna		18.°
Monaural pinna		23.°
Shoulder bounce		67.°

[1]Adapated from Searle, et al. (1976)

[2]The expected value of the standard deviation of localization judgements assuming only the corresponding cue(s) are available.

[3]The standard deviation of these three cues could not be estimated separately, based on the data used.

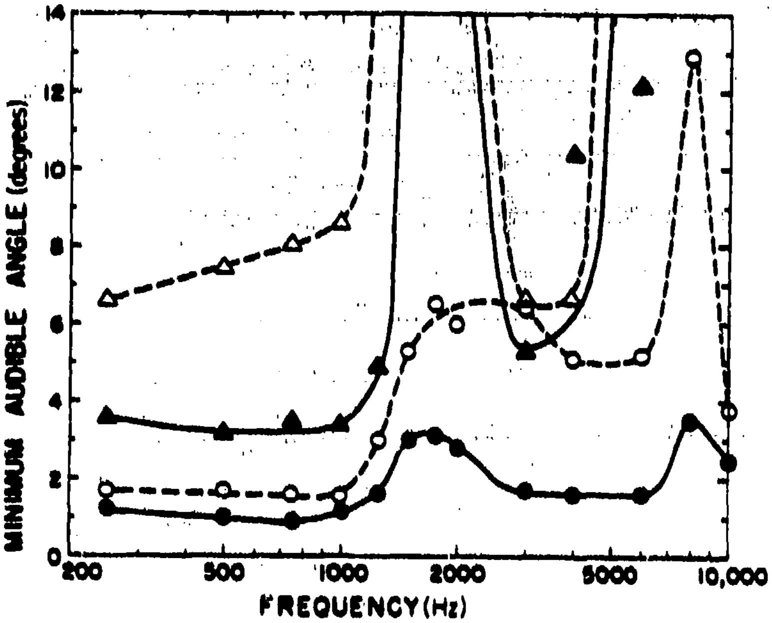


Figure 4. Minimum audible angle between successive pulses of tone as a function of the frequency of the tone and the direction of the source (•:0°; ∘:30°; ▲:60°; △:75°). (From Mills, 1972.)

in front of the listener. The MAA increases as the azimuth of the initial sound increases, to the point where it's indeterminately large at some azimuths and frequencies. The curves for the 0° and 30° conditions in Figure 4 exhibit two zones in which acuity deteriorates and then improves again with increasing frequency (from about 1200 to 2000 Hz and again in the region around 8 kHz). These zones may represent transitions from dependence on one type of cue to another. These results suggest that horizontal-plane localization is based on three or more types of acoustic cues.

In the following section is examined the available evidence for five of the six suspected cues for horizontal-plane localization. The review covers the physical characteristics of each cue, its function in terms of auditory localization performance, and what little is known about the interaction of cues. Discussion of the sixth cue, binaural pinna disparity, is covered in the Section 2.3, which concerns median-plane localization.

2.2.1 Interaural Time Differences

Depending on its angle, Θ, with respect to the listener's head, the sound from a distant source arrives at one ear before it arrives at the other. A simple geometrical model of the difference in the path that the sound traverses to one ear versus the other is shown in Figure 5. The path length difference is:

$$\Delta d = r\ (\Theta + \sin \Theta) \qquad (1)$$

Where Θ is measured in radians and r is the radius of the head, approximately 8.75 cm. The ITD for a sound at angle Θ is then:

$$\Delta t = \frac{\Delta d}{c} \qquad (2)$$

Where c is the speed of sound, about 343 m/s.

The ITD varies with the direction of the sound relative to the head as shown in Figure 6. The dotted lines represent two different sets of predictions. The lower dotted line is derived from equations (1) and (2). This line fits the measured ITD's for high frequency and broad-band sounds quite well. The upper dotted curve is derived from the equation:

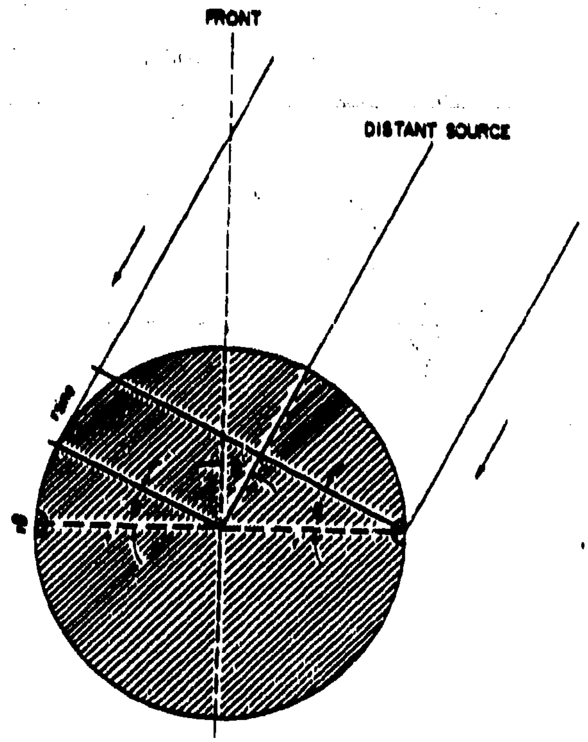


Figure 5. Differences between the distances of the ears from a source of sound far enough away to produce a nearly plane wave front. (From Mills, 1972.)

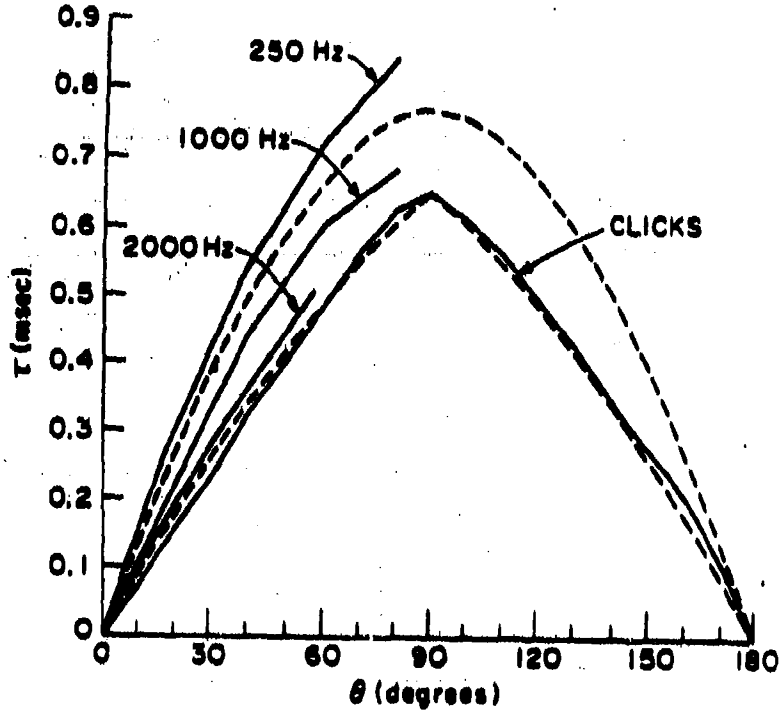


Figure 6. Interaural time delay for tones and clicks as a function of azimuth, θ. (From Durlach & Colburn, 1978.)

$$\Delta t' = \frac{\Delta d'}{c} = \frac{r}{c}(3 \sin \Theta) \qquad (3)$$

which is based on diffraction theory. It provides a rough approximation for low frequency sounds.

Kuhn (1977) developed a model of ongoing ITD's which provides a better prediction of ITD's for low frequency sounds. He compared the model predictions to measured ITD's on an anthropometrically-scaled manikin with pinnae. The results are shown in Figure 7. The ITD is greatest at low frequencies (< 500 Hz) and roughly frequency-independent below that value. The ITD is smaller, and again frequency-independent above about 3.0 kHz. For angles less than 60° from the median plane, ITD was smallest between 1.4 and 1.6 kHz (and just slightly less than the ITD at 3.0 kHz). Kuhn also noted that the rate of change of ITD with angle of incidence is greatest near the median plane (i.e., when the sound comes from in front of the listener). Therefore, the change in ITD with head rotation, if permitted, will be greatest when the source is located in the frontal region.

At low frequencies (< 1500 Hz), the ears are sensitive to both ongoing ITD's (or interaural phase differences - IPD's) and to transient ITD's. At high frequencies the ears are sensitive only to transient ITD's.

For any ITD, ongoing or transient, there are always two source positions which could have produced the ITD, one in front of the observer and the other behind. Note in Figure 8 that for source position B, the path lengths to the left and right ears are identical to those for source position A. This is an idealization since real heads are not circular and the ears are located more than 90° from the front of the median plane. For real observers, the back source position producing the same ITD will not be exactly 180° minus the front position. Listeners frequently make front-back reversals in judging sound source location, especially when pinna cues are not available (cf., Musicant & Butler, 1984a). Front-back ambiguities are discussed further in Section 2.2.3.

Additional ambiguities occur for ongoing ITDs produced by pure tones whenever the ITD is equal to or larger than one-half the period of the acoustic event. When this condition occurs, the listener will not be able to distinguish whether the source is located on the left side of the head, or at

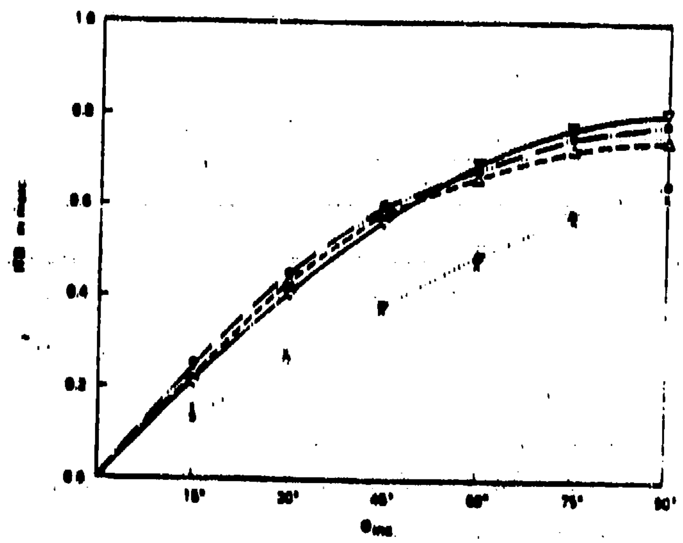


Figure 7. Measured and predicted interaural time delays with no torso as a function of angle of incidence • • •measurements at 360 Hz; △△△measurements at 500 Hz; ▽▽▽predictions based on Kuhn's (1977) model; ××× measurements at 3.0 kHz; ooopredictions based on simple geometry (See Figure 5).

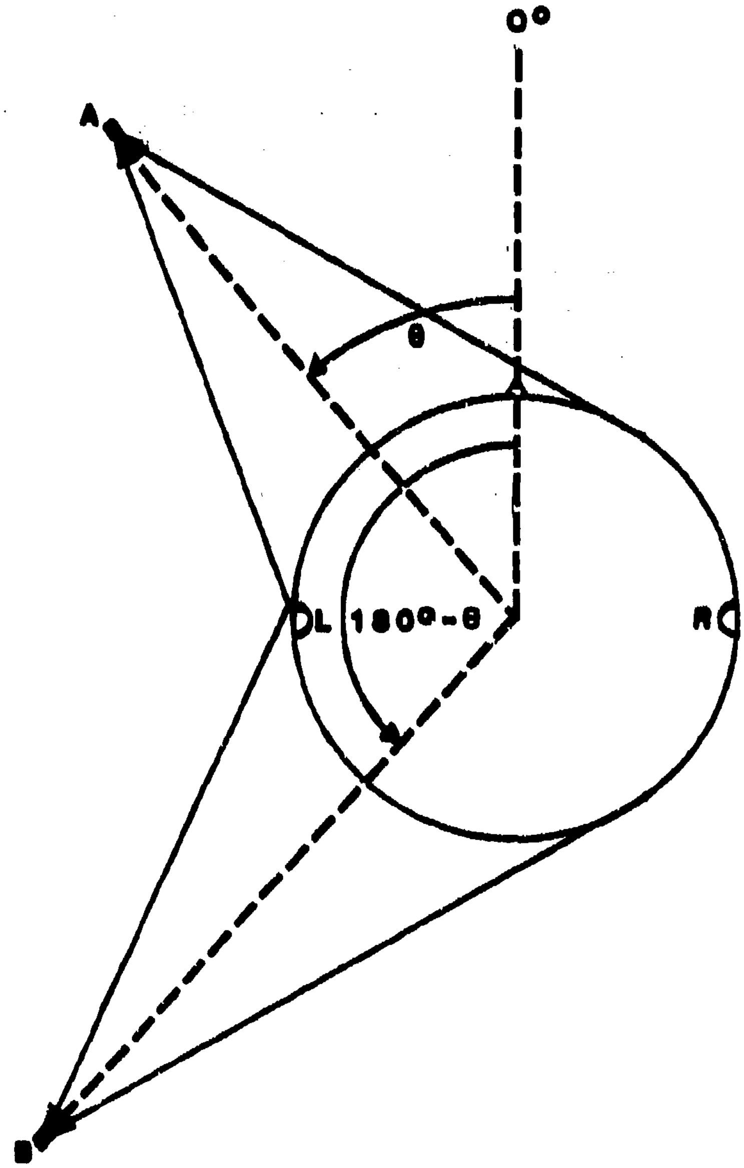


Figure 8. Front-back ambiguity of interaural time and amplitude differences.

the mirror image position on the right. Figure 9 shows a time line and represents the ITD as an interval on the line. Suppose, as shown, that a sound coming from the left side of the listener's head has a half-period equal to the ITD. Then, at a given moment in time, the left ear (L) is located at a pressure maximum when the right ear (R) is at a minimum. Now consider a sound coming from the right side of the head which is 180° out of phase with the sound coming from the left. It also produces a maximum at L when there is a minimum at R. The listener cannot perceive the absolute phase of a sound; hence he or she cannot distinguish such a sound coming from the left from a sound from the right.

The difference in the acoustic path length to the two ears is greatest when the sound approaches from $\theta = \pm 90°$. This condition therefore produces ambiguous ITD's starting at the lowest frequency of any direction of arrival. For such a sound, the shortest aerial path from ear to ear is about 23 cm. This distance exceeds half a wavelength (and therefore ITD exceeds half a period) at frequencies of about 750 Hz and greater.

Since the ITD decreases as the direction of the sound, θ, approaches 0° or 180°, higher frequency tones produce unambiguous ITD's in these regions. For example, at $\theta = \pm 45$, tones of frequency up to 1300 Hz produce an unambiguous ITD (the listener can distinguish whether the sound is coming from the left or right side of the head, but not necessarily from front versus back). Physiological evidence indicates that individual neurons of the auditory nerve can fire in synchrony with a periodic sound source only up to about 1000 Hz, due to the finite refractory period of the neuron. Above that, bundles of neurons can follow a periodic sound only up to about 1600 Hz (Tasaki, 1954). Therefore, no matter how small the ITD, the ear can only follow the interaural time (or phase) difference created by a periodic source up to about 1600 Hz.

Figure 10 shows the physical IPD's and IAD's produced by moving a source a just-noticeable-distance from the median plane. Also shown are the thresholds for IPD and IAD for tone pulses presented dichotically, via earphones. The interaural phase just-noticeable difference (IPjnd), or threshold, and the interaural amplitude just-noticeable difference (IAjnd) are greater than the physical IPD and IAD because the threshold (jnd) values of IPD and IAD were taken as those values that the listener correctly detected on 75% of the presentations (see Mills, 1972).

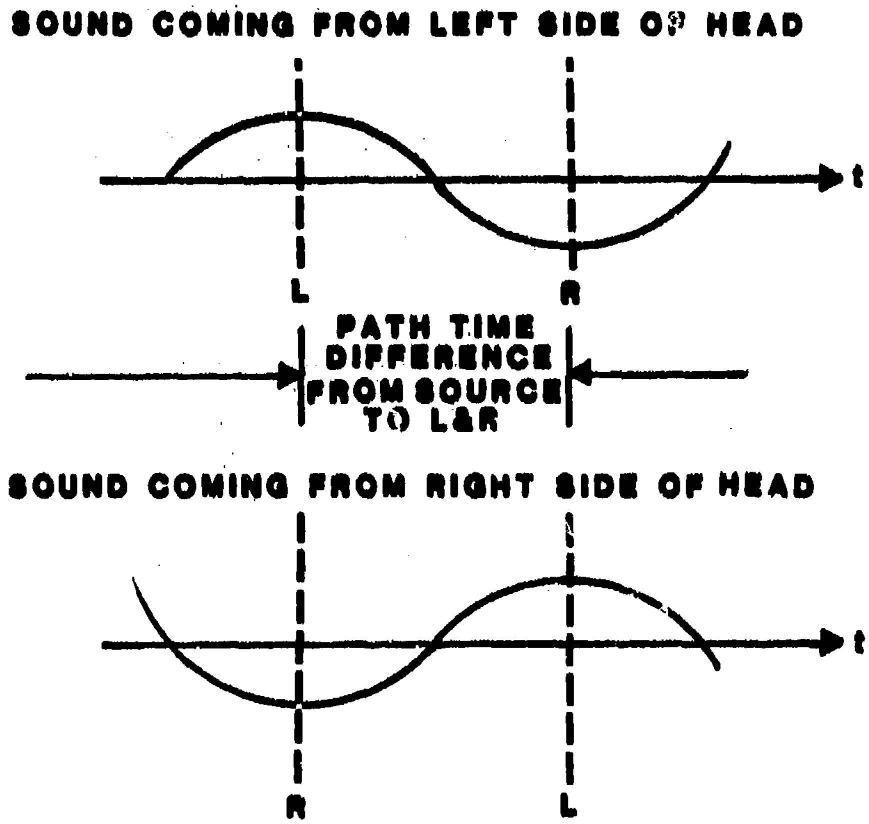


Figure 9. Momentary sound pressure level at the left (L) and right (R) ears. See text for explanation.

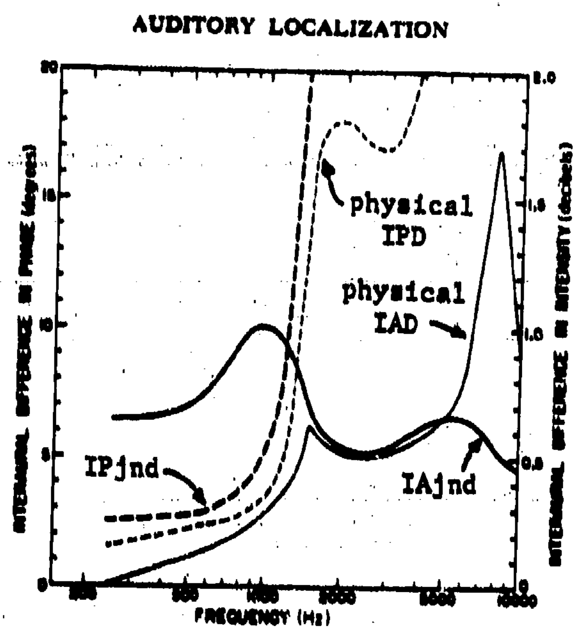


Figure 10. Comparison of the interaural differences in phase and intensity that can just be detected when tone pulses are presented through earphones with the interaural differences in phase and in intensity that are present when an actual source of tone pulses is moved just noticeably out of the median plane. (From Mills, 1972.)

The important feature of Figure 10 is that the IPjnd follows the physical IPD up to about 1200 Hz and then becomes indeterminately large. At about 1400 Hz, the IAjnd converges with the physical IAD and remains close to it up to about 6000 Hz. Above 6000 Hz, the physical IAD becomes much larger than the IAjnd. These data, along with the transition from 1200 Hz to about 2000 Hz seen previously in Figure 4, suggest that localization of pure tones is based on IPD up to about 1200-1500 Hz and on IAD above 1500 Hz. The data in Figure 10 provide no explanation for the transition seen around 8 kHz in Figure 4.

Although the auditory system is insensitive to ITD for pure tones at high frequency, it is sensitive to ITD for some complex high-frequency sounds. Sensitivity to ITD for high frequency sounds has been demonstrated when either of two conditions are met: (1) the sound has a low frequency periodic envelope, or (2) the sound includes a transient component.

A study by McFadden & Pasanen (1976) provides an example of the latter condition. Using dichotic headphone presentation, they measured the threshold for 80% correct lateralization of a two-tone complex. The tones were both in the 4000 Hz region, and differed in frequency by 25 to 550 Hz in different conditions. An interaurally uncorrelated, low-pass noise was used to mask the difference tone created by the mixing of the two primary tones. The complex tone was presented for 200 msec with 25 msec rise and decay times. Listeners detected ITD's as small as 27 μsec, based apparently on the transient (envelope) features of the stimulus.

The former condition for sensitivity to ITD's at high frequency is illustrated by a study by Yost, Wightman, and Green (1971). Listeners were asked to discriminate between high- and low-pass filtered clicks which were presented simultaneously (diotically) or at a time offset (dichotically). When a simple click was presented, high-pass clicks were more difficult to lateralize than were low-pass clicks. However, when the clicks were repeated 64 times during the observation interval, high- and low-pass clicks were lateralized equally well. Apparently, the auditory system is sensitive to the ongoing ITD produced by a low-frequency periodic envelope.

The detectability and usefulness of ITD's produced by transient sounds is not limited by the refractory period of the neurons or phase ambiguities. Hence transient ITD's are effective even at frequencies as high as 5000 Hz (cf., Yost et al., 1971).

2.2.2 Interaural Amplitude Differences

The interaural amplitude difference produced by the head shadow varies as a function of the frequency of the sound as well as its direction. Figure 11 shows the IAD measurements for pure tones on five human subjects. As expected, the IAD is near zero at 0° azimuth, and reaches a maximum between 60° and 120°. At frequencies less than 1800 Hz, the IAD is relatively small for all angles of incidence. At 2500 Hz the IAD reaches a maximum of about 12 db near 80°. As the frequency increases, the IAD approaches 20 db at some angles. These data are in agreement with Figure 10, which shows that the IAjnd and the actual IAD produced by moving a source near the median plane follow one-another closely from near 1500 Hz to about 6000 Hz.

Another important feature of the IAD curves in Figure 11 is that the rate of change of IAD with azimuth is greatest near 0° and generally flattens out as the source moves to the side of the head. Therefore, any rotation of the sound source or the head in the horizontal plane should produce a much greater effect near 0° than at larger angles. This finding agrees well with Figure 4, which shows that the minimum audible angle increases rapidly as the azimuth of the source increases.

As noted earlier in the discussion of Figure 10, at frequencies greater than 6000 Hz, the close correspondence between IAjnd and the physical IAD ends. Blauert (1983) found that IAD changes erratically as a function of frequency above 6000 Hz. Other evidence (see section 2.2.3) suggests that spectral changes produced by the pinnae may be the dominant cues at the higher frequencies.

2.2.3 Monaural Cues

Interaural time and intensity differences enable the listener to discriminate sounds coming from the right vs. left side of the head, but do not differentiate sounds in front from their mirror-image position in back of the head. There is now evidence that resonances, reflection, and diffraction from the pinnae, torso, and head serve as monaural cues which enable the listener to make front-back discriminations. These anatomical features produce direction-dependent changes in the spectrum of the sound received at the eardrums.

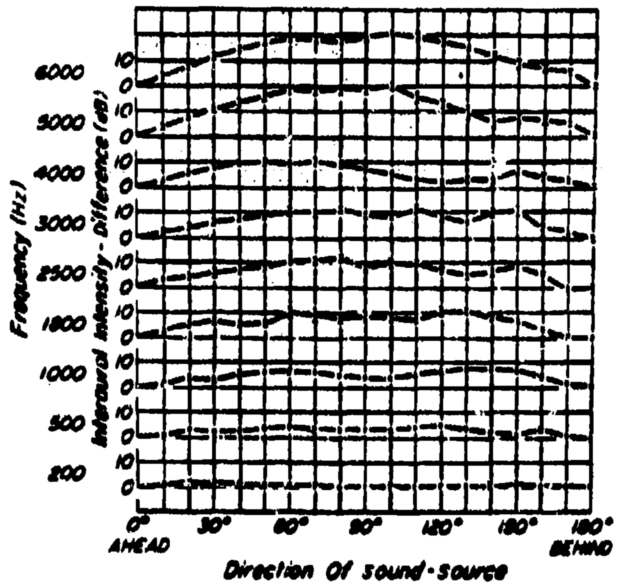


Figure 11. Interaural intensity differences as a function of the direction of the source and the frequency of the sound. (From Feddersen Sandel, Teas, and Jeffress, 1957.)

In order for monaural cues to be effective aids to localization, the listener must have a priori information about the spectrum generated by the source. Only then can the listener judge whetner a given spectral feature is a function of the source position, or a feature of the free-field source spectrum. Such a priori information can be gained through pre-experimental knowledge or repeated presentations of the source.

The monaural cues seem to consist of relative amplification of one region of the spectrum relative to another region. If the source spectrum is limited to a very narrow bandwidth, as it is for pure tones, then little or no relative amplification is possible. If, further, the amplitude of the source is varied from presentation to presentation, any noticeable attenuation or amplification of the sound from one exposure to the next is no longer a useable cue for localization.

Stevens and Newman (1936) found that tone bursts coming from in front of the head were often perceived as coming from the back. When the intensity of the sound was also varied randomly from one presentation to the next, listeners performed at nearly chance level. On the other hand, Musicant & Butler (1984a) found that wide-bandwidth noise stimuli were localized with very few front-back reversal errors when they contained significant energy from 1 to 4 kHz and greater. However, when the listener's pinnae were occluded in the same study, the percentage of front-back reverals increased dramatically for both wide-band noise and for 4 kHz high-pass noise, but not for 4 kHz low-pass noise. These results strongly suggest that the pinnae produce effective localization cues at frequencies of 4 kHz and greater.

In the Musicant and Butler study, when the pinnae were not occluded, the 4 kHz low-pass noise, the 4 kHz high-pass noise and the wide-band noise all produced very few front-back reversals, whereas the 1 kHz low-pass noise produced front-back reversals on about 50 percent of the trials. This suggests that some cues also operate in the region from 1 to 4 kHz to reduce front-back reversals. When the pinnae were occluded, the number of front-back reversals increased in the wide-band and 4 kHz high-pass noise conditions, but not for the 4 kHz low-pass noise condition. This result suggests that the pinnae produce cues only above 4 kHz, and that other anatomically-related cues (such as torso reflection and monaural head shadow) operate in the 1 to 4 kHz band to help resolve front-back ambiguities. Gardner (1973) reported acoustic data (HRTF's) for manikins with and without a torso. The torso clearly produced amplitude variations in the region from 0.7 to 3.5 kHz.

As mentioned earlier, a number of acoustic investigations have been done to directly measure the spectral changes produced by the pinna, head, and torso (Gardner, 1973; Hebrank & Wright, 1974b; Mehrgardt & Mellert, 1977; Rodgers, 1981; Shaw, 1974a; 1974b; Shaw & Teranishi, 1968; Weinreich, 1982). The resulting HRTF's show the ratio of sound pressure amplitude (and/or difference in phase) at the ear drum or ear canal entrance versus the amplitude or phase in the free field (when the listener is absent). As noted earlier, there are numerous large differences in amplitude response at particular frequencies for sound sources in front versus the mirror-image position in back which could serve as localization cues. Shaw (1974b) presents amplitude HRTF's normalized to 0° azimuth, which he calls "azimuthal dependence" curves (see Figure 12). Careful inspection of these curves suggests possible acoustic features which could play a role in front-back discrimination and monaural localization cues generally. He notes that the overall gain (relative to 0°) increases smoothly from -45° to +45° azimuth. From 60° to 120° the gain remains about constant, except in the 2-6 kHz region where it decreases smoothly. From 120° to 165° there is a fairly constant increase in gain from 2-5 kHz accompanied by a steady decrease at other frequencies. Shaw concludes that these regular changes in amplitude response are: (1) due to the pinna, and (2) provide the physical basis for (monaural) localization. Further psychoacoustic research is needed to determine the necessary and sufficient acoustic features for monaural localization.

Most investigators have reported HRTF's averaged over a number of listeners. These averaged HRTF's show broad, common trends in amplitude response, but the averaging process eliminates much of the "fine structure" of individual HRTF's. Mehrgardt and Mellert (1977) shifted their HRTF's along the frequency dimension before averaging in order to preserve major features common to individual HRTF's. Rodgers (1981) has suggested that ideosyncratic features of an individual's HRTF may be important for localization performance.

A number of studies show that listeners quickly adapt to a change in the ideosyncratic features of the HRTF. When SAL cues derived from artifical pinnae, or models of someone else's pinnae, are presented to an observer via

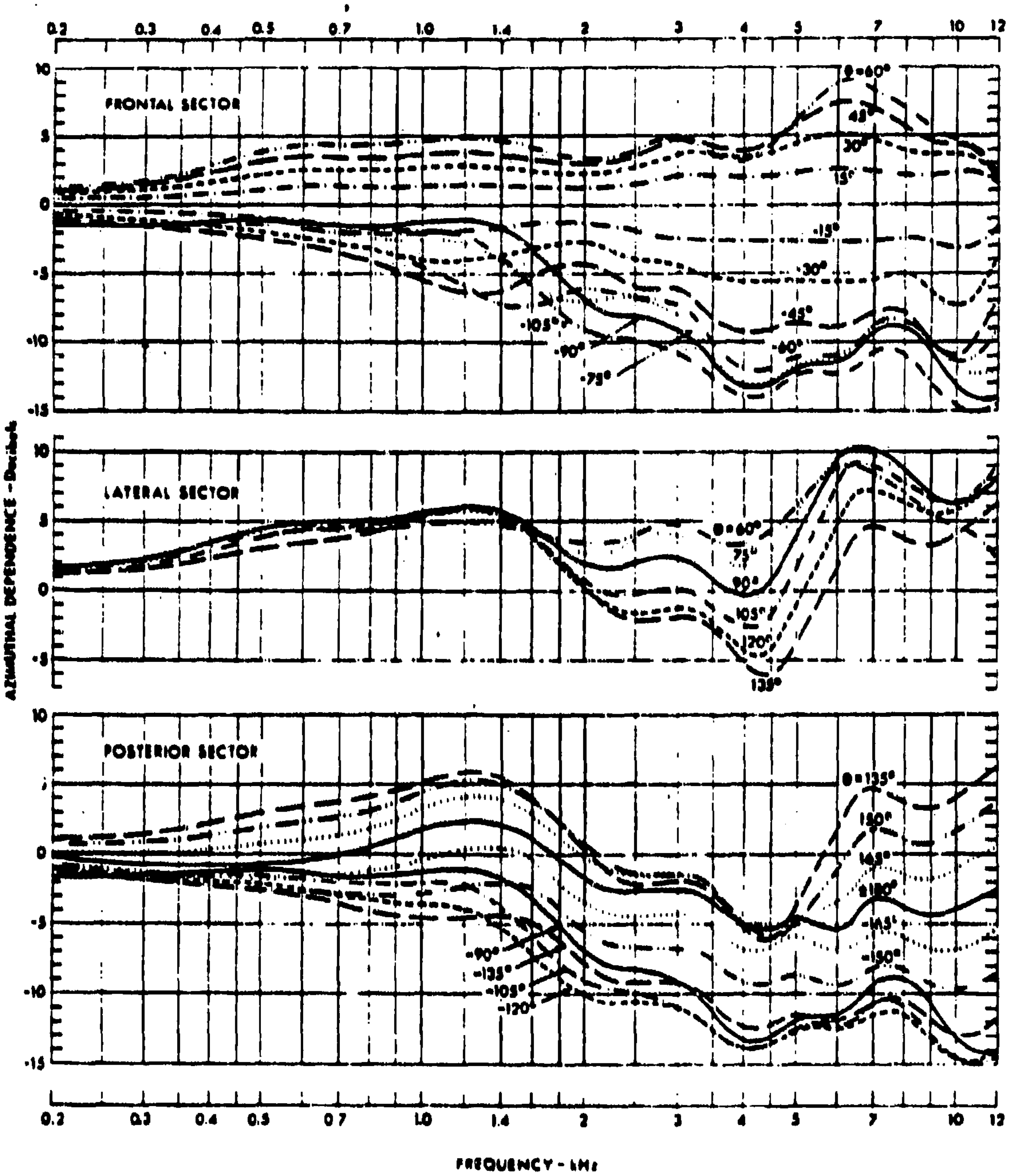


Figure 12. Family of 23 self-consistent curves showing HRTF's normalized to 0 degrees azimuth. (From Shaw, 1974b.)

earphones, localization is initially less accurate than with one's own pinnae. However, adaptation to the foreign pinnae occurs rapidly. After several minutes of practice, most observers are able to localize nearly as accurately as with their own pinnae (cf., Searle, 1982; Batteau, 1967).

It also appears that the anatomical features of some listeners provide better localization cues than do those of other listeners. Butler and Belendiuk (1977) reported that some listeners localize more accurately when presented with cues generated from another person's pinnae than with their own.

The acoustical studies reviewed above relate the direction of a sound source to the proximal acoustic stimulus. While this is necessary and useful research, it is not a sufficient account of auditory localization performance. Psychoacoustical studies are needed to relate the proximal acoustic stimulus to the listener's perception of the source location.

One of the problems encountered in relating the proximal stimulus to the percept in auditory localization has been separating the effects of the various cues, such as ITD's, IAD's, pinna cues, etc. One partially successful approach has been to control the situation under study by using pure tones as stimuli. This work, reviewed previously, has helped to identify the roles of ITD's and IAD's. However, this approach has not been successful in isolating pinna, torso, and monaural head cues because IAD's still play a role, even at high frequencies.

A more productive approach has been to study monaural localization, thereby eliminating the binaural cues such as ITD and IAD. Butler and his colleagues have pursued this approach for horizontal as well as median-plane localization. Their results suggest that source angular position is frequency-coded, much like Blauert's (1969/70) "critical band" hypothesis for median plane localization.

In a series of studies (Butler and Flannery, 1980; Flannery and Butler, 1981; Musicant and Butler, 1984b), Butler and his co-workers found that the perceived direction of narrow-band, high frequency noise is related to its center frequency (CF) rather than its actual direction. In these experiments, the right ear was acoustically blocked and the stimuli were presented randomly from various azimuths in the front-left quadrant or whole left side. A 1 kHz-wide band-pass noise centered at 4 kHz was nearly always reported as coming from directly in front of the listener, regardless of its actual position. As the CF of the noise was increased, the perceived location tended to move to the side, generally reaching 90° when the CF reached about 8 kHz. As the CF increased beyond 8 kHz, the apparent location moved back directly in front for most listeners, and then migrated laterally as the CF was increased further. When the CF reached 12 to 13 kHz, the apparent location again moved back in front of the listener. Not all listeners followed this pattern, and there were considerable individual differences among those who did follow the pattern as to what CF's were associated with which directions. For a few listeners, the apparent location of the sound moved from near 0° azimuth through the rear quadrant to nearly 180° as the CF increased from 4 to 12 kHz.

Flannery and Butler, and Musicant and Butler, measured the relative gain for the free-field versus the entrance to subject's ear canals for 1 kHz-wide noise stimuli with various CF's. They found that the gain for any given stimulus tended to be greatest at that azimuth that listeners most often perceived as the source location. However, the amplitude differences between locations for given stimuli were frequently as small as 1 db. Similar trends can be seen in the HRTF's reported by Shaw (1974b) and Mehrgardt and Mellert (1972). In Shaw's Figures 6 and 7, the HRTF for a 4 kHz tone shows a broad peak at about 40° azimuth. As the frequency increases to 8 kHz, the peak migrates to 90° or slightly greater azimuth. A similar trend can be seen in Mehrgardt and Mellert's Figure 18. Thus there is at least a partial correspondence between the features of averaged HRTF's and the perceived source location in the horizontal plane.

Butler and his colleagues call each frequency region which moves the sound from the front to the lateral position a Spatial Reference Map (SRM). For example, for most listeners, 4-8 kHz is an SRM, as is 8-12 kHz. Butler and Flannery and Flannery and Butler compared monaural localization accuracy for 4 kHz-wide noise with various CF's. In one condition, the CF was such that the noise bandwidth spanned two SRMs, in another condition the noise bandwidth fell entirely within an SRM. Localization performance was better when the noise-bandwith was centered over an SRM boundary. It is not clear why this should be so, and the authors did not offer a satisfactory explanation.

Weinrich (1982) reported strong evidence for front-back discrimination cues around 1.2 kHz and from about 3.5 to 6 kHz. He measured the HRTF from the free field to near the ear canal entrance for azimuths of 30°, 150°, 210°, and 330° for four listeners. He then subtracted the amplitude response generated by a source in back from that generated by a source in the corresponding front position (e.g. 150° response minus 30° response). This was done for both the ipsilateral and contralateral ears. In the ipsilateral ear, he found a broad, marked increase in SPL from about 3.5 to 6 kHz and a small decrease around 1.2 kHz. In the contralateral ear, he found a marked decrease around 1.2 kHz. He next addressed the question of whether these front-back acoustic differences are actually used as cues for front-back discrimination. Using the HRTF features noted above, he produced simulated front and back cues. The cues were impressed on speech and white noise stimuli and delivered via earphones. Appropriate ITD's were also introduced to simulate front and back positions at various azimuths. Subjects were able to assign these synthesized signals to an array of 12 directions with relatively few front-back reversals. However, the spectral features impressed on the signal had to be exaggerated in order to achieve good front-back discrimination performance. The partial success achieved by Weinrich suggests that it may be possible to produce accurate SAL performance based on variations in the spectrum as a function of apparent azimuth.

2.3 Localization in the Median and Vertical Planes

The major binaural cues, interaural time difference (ITD) and interaural amplitude difference (IAD), produce not only front-back but also above-below localization ambiguities. For the idealized spherical head, the ITD and IAD specify the location of a sound source to within a "cone of confusion," shown in Figure 3. Any source location on the cone produces the same ITD and IAD. Of course real heads are not spherical, but this does not eliminate the ambiguities. It simply makes the cone an irregular rather than smooth surface.

Localization in the median sagittal plane (MSP) is of interest because ITD's and IAD's are presumably absent there. Many investigators have assumed that localization in the MSP is based solely on cues produced by the pinnae and torso. In contrast, for judgements of source elevation in other vertical planes, ITD and IAD cues should be available, as well as pinna and torso cues. One would therefore expect judgements in the MSP to be considerably less accurate than comparable judgements in the other vertical planes. Gardner and Gardner (1973) measured localization accuracy for connected speech. Loudspeakers were arranged in a semicircular arch over the subject's head, spanning from 0° elevation in front of the subject to 180° in back. For different test conditions, the subject was rotated ±5°, 15°, 45°, or 90° with respect to the arch in order to test localization in other vertical planes. The results, shown in Figure 13, show that elevation judgements are most accurate for the transverse plane (90° to the median plane) and intermediate for oblique planes. These results make sense in terms of the expected change in ITD and IAD produced by a given change in elevation, $\Delta\epsilon$. In the median plane, $\Delta\epsilon$ produces zero change in ITD and IAD; in the transverse plane the change is maximized; in oblique planes the change in ITD/IAD with $\Delta\epsilon$ is intermediate.

Localization accuracy also varies considerably within the MSP. Wettschureck (1973) measured the minimum audible angle for a white noise source at various locations in the MSP. He found the greatest acuity, about 4° in the front of the listener, intermediate acuity behind, and least acuity (about 10°) overhead. This compares with a maximum acuity of 1° in the frontal horizontal plane.

2.3.1 Interaural Differences

Studying localization within the MSP would seem to be a convenient way of eliminating ITD and IAD, therefore enabling the investigator to study torso and pinna cues in isolation. However, the assumption that interaural differences play no role in MSP localization has been questioned.

As shown in Table 1, Searle et al. (1976) propose a third interaural cue that provides information about the elevation as well as the azimuth of the sound source. This third cue is the interaural pinna amplitude difference. Asymmetries in the left and right pinnae (and head asymmetries) could produce interaural amplitude differences as a function of source elevation in the MSP.

Searle, Braida, Cuddy & Davis (1975) measured the amplitude response of the head and pinnae at the ear canal entrance as a function of frequency. The measurements were taken on three human subjects, for sources at various

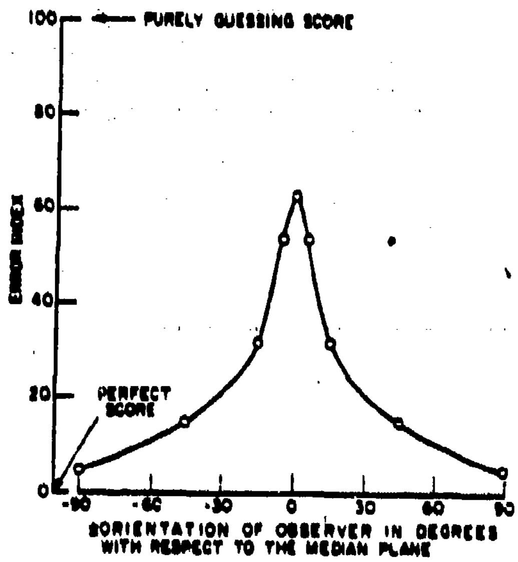


Figure 13. Localization of connected-speech signals as function of transverse to median plane orientation. (From Gardner and Gardner, 1973.)

elevations in the MSP. As expected, there was a common change in amplitude response over frequency for both ears. These are the monaural pinna cues. In addition, Searle et. al. found a disparity between the left and right ear responses which changed systematically with elevation angle. These differences are potential interaural cues to localization in the MSP.

Searle et al. (1975) argue that the interaural pinnae differences are used as cues for MSP localization. They review previous research which shows that monaural localization in the MSP is less accurate than binaural localization in the MSP (cf., Butler, 1969a). This finding could be due to either the absence of binaural disparity cues, or to the loss of redundant information from one ear in the monaural condition.

As a test of these two alternate explanations, Searle et al. (1975) presented simulated localization cues for sources in the MSP via earphones to four subjects. The localization cues were obtained by first recording from the left and right ear canals of each subject for sources in the MSP. The recordings were then played back either dichotically (left channel to left ear and right channel to right ear) or diotically (left or right channel to both ears). Searle et al. also varied the amount of _a priori_ information about the signal available to the subjects by varying the spectrum of the signal from trial to trial in some test conditions. As noted in Section 2.2.3, the listener must be familiar with the source spectrum in order to use monaural cues effectively. They reasoned that reducing the amount of _a priori_ information available should affect localization accuracy less in the dichotic than in the diotic condition if binaural disparities were used as cues in the former. Their results show better localization in the dichotic conditions overall and greater degradation of localization accuracy in the diotic condition when the source spectrum was varied from trial to trial.

The Searle et al. (1975) results suggest that binaural disparity cues are used in MSP localization, but they are not conclusive. In the diotic presentation condition, one of the ears received simulated cues generated by recording from the opposite ear. In the binaural condition, both ears received cues generated by recording from themselves. It is possible that the combination of unnatural cues in one ear plus the experimental manipulation of trial-to-trial spectral variations in the source produce an interactive effect on _monaural_ localization accuracy. For example, if Searle et al. had also tested their subjects monaurally, one might find that varying the source spectrum produces greater degradation when the cues were recorded from another ear, than when the cues were recorded from the ear being stimulated.

On the other hand, Hebrank and Wright (1974a) argue that binaural cues do not play a role in MSP localization. In a first experiment, they presented white noise and "rippled-spectrum" noise to monaural and binaural subjects, with the source in the MSP. The spectrum of the rippled noise contained peaks and notches similar to those introduced by the pinna. The location of these peaks and notches was varied randomly from trial to trial. The rippled noise presumably denied to the subjects _a priori_ information about the signal spectrum, and should therefore degrade localization based on monaural cues. Hebrank and Wright found that the rippled noise (versus white noise) produced the same amount of degradation in localization performance for monaural and binaural subjects. This result suggests that binaural information, even if present in MSP localization, does not aid localization. In a second experiment, Hebrank and Wright found that, if given localization feedback, monaural subjects quickly learned to localize in the MSP as accurately as do binaural subjects.

Even though binaural differences may play no direct role in MSP localization, they are responsible for centering the acoustic image in the MSP. As Blauert (1982) points out, binaural information is not absent in MSP localization. Rather, binaural differences are at that value which indicates that the source is located in the MSP. Gardner (1973) also points out that binaural information performs a "centralizing" function in MSP localization. When one ear is blocked, the acoustic image rotates laterally toward the open ear. In this condition, judging the position of the source in the MSP becomes very difficult.

2.3.2 Monaural Cues

The two dominant theories of monaural localization are the theory of Timbre Differences (a frequency domain approach) and the Echo-Delay position (a time domain approach). The latter position was argued by Batteau (1967) to explain how the pinnae produce direction-dependent cues. Specifically, he proposed that reflections from the various folds and cavities of the pinnae produce reflected, and therefore delayed, signals. The relative amount of each delay is postulated to depend on the location of the sound source vertically and horizontally relative to the pinna. Batteau further proposed that the ear and brain determine the sound source location by measuring the amount by which the echoes are delayed from the main signal.

The neurological delay-processing model proposed by Batteau has been largely discarded due to claims that the temporal resolution of the auditory system is not adequate. The minimum audible angle in the median plane (where processing is based only on monaural cues) is about 4° (Harris & Sargeant, 1971; Wettschureck, 1973). In order to discriminate that small a change in source location, the temporal acuity of the ear and brain would have to be less than 5 μsec (Batteau, 1967). Green (1971) and Zwicker (1973) report that the minimum discriminable monaural time difference is greater than 1 msec. However, Hebrank and Wright (1974a) found that the just-noticeable-difference (jnd) for time delays was in the region from 5 to 7 μsec. They used time-delayed white noise summed with itself and presented monaurally. Hebrank and Wright conclude that the human auditory system has the required temporal acuity, but reject the neurological delay-processing model on other grounds. They base their rejection on a report by Thurlow and Runge (1967), who found that clicks were localized more poorly than was white noise. If the auditory system determined location by judging delays, stimuli with steep rise times, such as clicks, should be localized more accurately than the relatively slowly varying envelope of a noise function.

The spectral approach to monaural localization holds that reflections, resonances, and diffraction of the sound source by the pinna and torso produce direction-dependent changes in the spectrum of the sound reaching the eardrum. The auditory system is presumed to judge the source location according to the spectral features of the sound reaching the eardrum. As in the case of monaural localization in the horizontal plane, _a priori_ information about the source spectrum is required in order to judge its location. Otherwise, the auditory system has no way of discerning whether a given spectral feature is a function of the source location or a property of the original source spectrum. The spectral theory leads naturally to the question of what spectral features are produced by the pinnae and torso and used by the brain.

A number of investigators have studied the frequency range over which the pinnae and torso produce useful cues for MSP localization. Gardner (1973) found that localization of full-band noise in the anterior portion of the MSP was degraded when the pinnae were occluded. He also compared localization accuracy for one-half octave bandwidth noise centered at 2, 3, 4, 6, 8, or 10 kHz, with and without the pinnae occluded. With open pinnae, localization accuracy increased as the center frequency of the noise-band increased; and was best for full-band noise. When both pinnae were occluded, localization accuracy dropped to near chance for the 4, 6, 8, and 10 kHz noise bands, and to just above chance for the 2 and 3 kHz noise bands and the full-band noise. These results suggest that the pinnae produce useful cues from 3 or 4 KHz to at least 10 kHz, and that some other factor (presumably the torso) produces cues in the 2-3 kHz region. Gardner and Gardner (1973) report comparable results for the posterior portion of the MSP.

Hebrank and Wright (1974b) measured localization accuracy for white noise and sharply filtered high- and low-pass noises. Source position was varied among nine locations from -30 to +210° in the MSP. Their results, presented in Figure 14, show that the absence of spectral energy below about 3.8 kHz and above 16 kHz does not affect MSP localization performance. These results are not in conflict with Gardner's results showing limited ability to localize sound in the 2-3 kHz region. Note in Figure 14 that localization accuracy is still slightly above chance for the 4 kHz low-pass noise.

The most compelling evidence for the spectral theory of monaural localization is the auditory illusion produced by narrow-band signals in the MSP. The perceived location of such sources is related to their center frequency rather than their actual location. Roffler and Butler (1968) presented tone bursts of frequencies from 250 Hz to 7200 Hz at locations from -13° to +20° elevation in the MSP. They found that reported source height on a 54 inch high panel in front of the listener was monotonically related to stimulus frequency. They also report several experiments which show that this relationship is not due to learned associations, e.g., the convention of labeling tones of greater frequency "higher".

Blauert (1969/70) presented 1/3-octave band-pass noise to binaural listeners from loudspeakers located in front, above, or behind the subject. The listener reported the apparent direction of the sound source (in front,

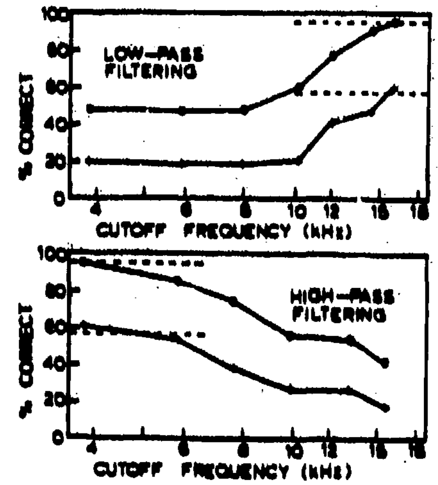


Figure 14. Mean percent correct responses for ten subjects localizing low- and high-pass-filtered white noise with accuracies of $\pm 45°$ (0) and $\pm 15°$ (+). Dotted lines show mean accuracies for white noise. (From Hebrank and Wright, 1974b.)

behind, or above). The center frequency, duration and intensity of the stimuli were varied randomly from trial to trial. Reported source direction was related to signal frequency, and not to signal duration, location, or intensity. Specific frequency bands were associated with each direction, as shown in Figure 15.

Butler and Helwig (1983) report results strikingly similar to those obtained by Blauert at frequencies of 4 kHz and greater. They presented 1.0 kHz-wide band-pass noise bursts at center frequencies from 4 to 14 kHz. The subject was seated under an arched array of loudspeakers which spanned from -30° to +210° elevation in the MSP. Reported source location was related to center frequency and not to actual source location. For four out of five listeners, reported source location increased nearly linearly from -15° to about +180° elevation as the center frequency increased from 4 kHz to 11 or 12 kHz. For one listener, reported source elevation increased to roughly overhead (90° to 105°) at 8 kHz, and then abruptly reverted to the frontal position at 9 kHz. For four of the five listeners, the spatial correlates of frequencies above 4 kHz in Butler and Helwig's study corresponded very closely to those in Blauert's study. Hebrank and Wright (1974b) also report apparent source positions for filtered noise which are consistent with the findings of Butler and Helwig, and Blauert.

Given such consistent psychoacoustic data, one would expect to find reliable corresponding spectral features. Hebrank and Wright (1974b) made models of three subjects ears and measured their amplitude response (HRTF) from 4 to 16 kHz for sources at various elevations in the MSP. They found three features of the HRTF's that were common among the three ears, and which corresponded to the perceived locations of filtered-noise stimuli: (1) a notch, the low-frequency side of which migrated from about 5 to 8 kHz as the source elevation increased from -30° to +60°, (2) a peak between 7-9 kHz for overhead source locations, and (3) greater energy above 13 kHz for frontal source locations than for behind locations. Butler and Belendiuk (1977) measured the amplitude response from 4 to 9 kHz for 8 subjects' ears. They reported a prominent notch, the center of which moved from about 5.5 kHz to about 7 kHz as the source elevation changed from -30° to +30°. The same features of the pinna amplitude response have also been reported by Shaw and Teranishi (1968) for sources at elevations of -45° to +45° in the transverse vertical plane.

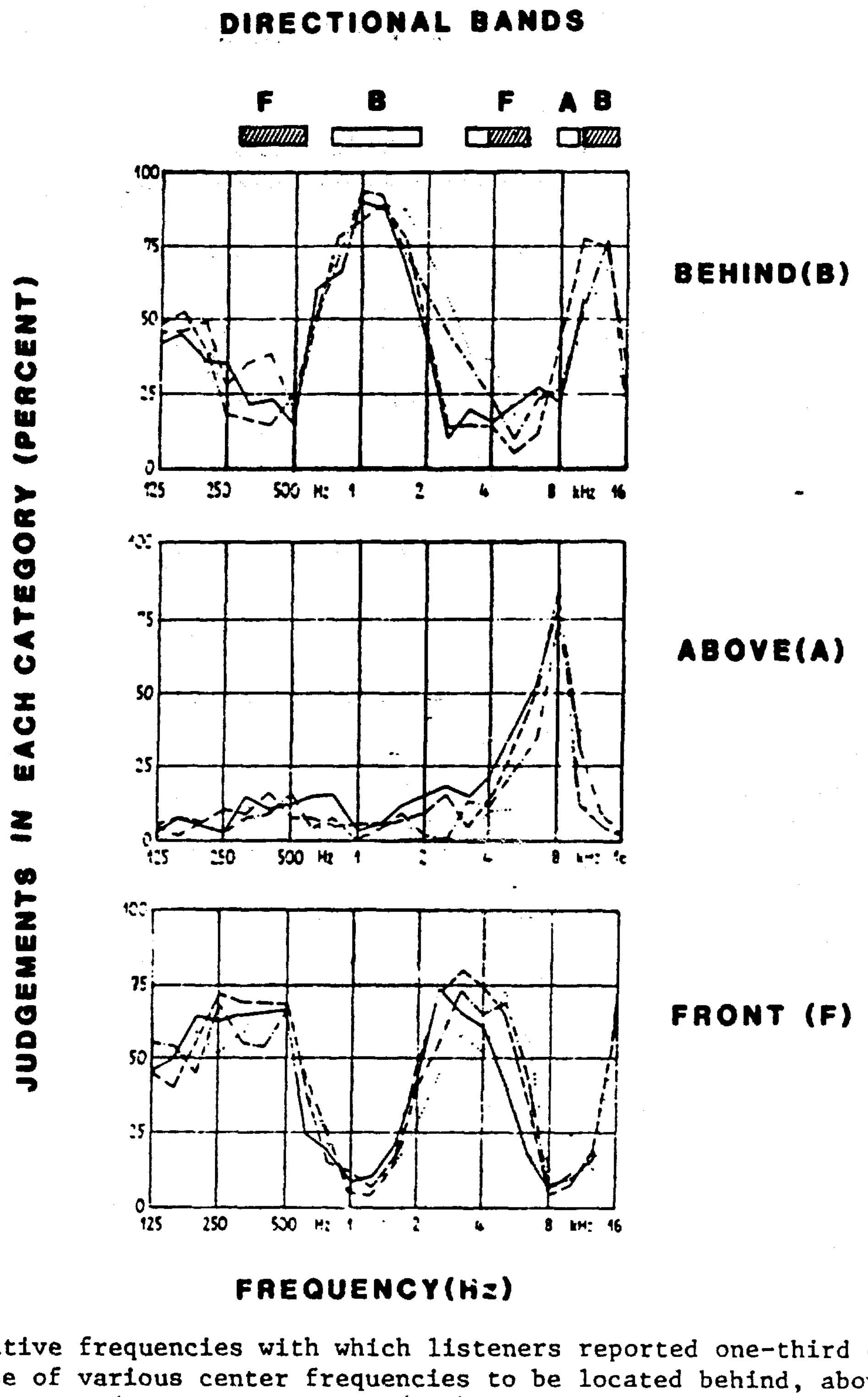


Figure 15. Relative frequencies with which listeners reported one-third octave noise of various center frequencies to be located behind, above, and in front. (From Blauert, 1969/70.)

A number of investigators have proposed that elevation in the MSP is encoded by a <u>combination</u> of spectral features. Wright, Hebrank, and Wilson (1974) point out that the spectral features used by the auditory system in monaural localization could be generated by pinna reflections added to the original signal. A broad-band signal when summed with a delay of itself at an appropriate amplitude ratio, α, produces a sound, the spectrum of which contains multiple peaks and notches as shown in Figure 16. The delays proposed by Batteau (1967) (10-80 μsec for azimuth angles and 100-300 μsec for elevation angles) produce ripples in roughly the appropriate parts of the spectrum. Watkins (1978) shows that the spectral cues which produce the perception of particular source directions can be described as components of a "comb-filter" pattern which can be produced by echo delays in the range of values suggested by Batteau. Figure 17 shows the effect of adding two delayed components of a white noise signal to itself. Note that a given delay, τ_v, produces a spectrum with multiple peaks and notches. The frequencies of the peaks and notches migrate as the amount of delay, τ_v, of one of the added signals changes.

Watkins conducted an experiment in which subjects reported the last perceived location of an apparently moving source. Apparent movement was created by varying one of the delays, τ_v, in a two-delay-and-add signal. The "above" and "below" labels in Figure 17 are positioned above the τ_v values which were perceived as above (+40°) and below (-20°) in the transverse vertical plane.

Note that the spectral features shown in Figure 17 correspond roughly to those reported by Hebrank and Wright (1974b), and Butler and Belendiuk (1977), for source elevations in the MSP. Specifically: (1) there is a notch, the low-frequency side of which migrates from about 5 to 8 kHz as the source elevation increases, and (2) there is a peak between 7-9 kHz for overhead (above) locations. Watkins concludes that: (1) the same monaural mechanism contributes to vertical localization in both the transverse and median planes, and (2) the decoding mechanism for vertical location is based on the recognition of spectral patterns like those produced by a multiple delay-and-add system.

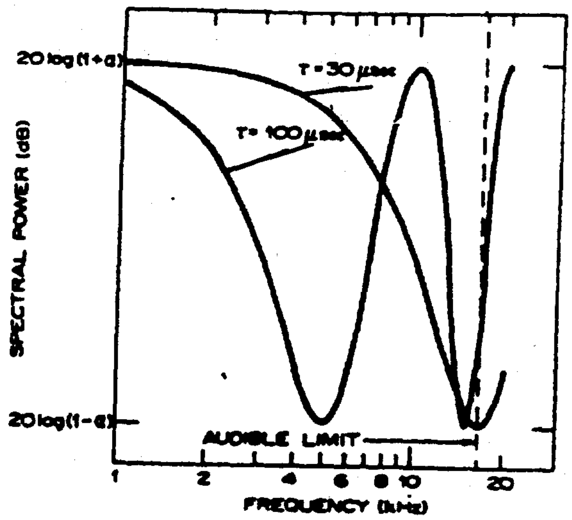


Figure 16. Spectral filtering created by combining a signal with a delay of itself. Delays less than 30 μsec cause low-pass filtering and delays greater than 30 μsec cause multiple spectral notches. (From Wright, Hebrank, and Wilson, 1974.)

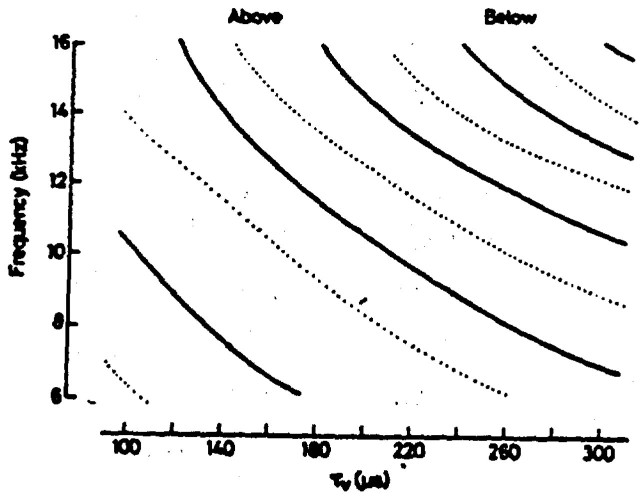


Figure 17. Spectral peaks and notches generated by adding white noise to two delayed echos of itself, as a function of one of the delays, τ_v. The "above" and "below" labels are positioned over τ_v values which listeners reported the sound source to be high or low in the transverse vertical plane. (From Watkins, 1978.)

2.4 Auditory Distance Perception

The acoustical characteristics of a sound field that vary as a function of the distance of the source have been well documented. However, the effectiveness of these acoustical parameters as psychophysical cues for the perception of distance has not been thoroughly studied. A distinction must be made between absolute judgements of distance and relative judgements of distance: absolute judgements can be made based exclusively on the characteristics of the sound field, whereas relative judgements require that the listener have, or assume, *a priori* information such as a reference sound or repeated exposure to the sound source. Monaural cues (such as the intensity and spectral composition of the sound at the eardrum) can provide only relative information. In principle, binaural cues can provide absolute information only if the direction of the source is known or assumed, or if the sound is repeatedly sampled as the head is rotated and the source is known to be stationary. There is considerable evidence that monaural cues play an important role in relative distance judgements. There is also evidence that binaural information plays a role in distance judgement when the head is stationary. However, it is not clear what binaural cues are used.

The physical characteristics of a sound field vary in terms of both intensity and frequency composition as a function of distance. According to the $1/R^2$ loss law, intensity decreases as distance increases. Amplitude decreases by 6 db for each doubling of distance. Coleman (1963), in his review of auditory depth cues, cites various psychophysical data which support the use of the $1/R^2$ loss law as a relative cue for distance judgements.

The usefulness of intensity as a cue to distance seems to depend on (1) the distance of the source from the listener, (2) the characteristics of the sound, and (3) the listener's degree of familiarity with the sound. Bekesy (1949) showed that the perception of distance tends to level off as distance increases past a critical point. This outer limit on distance perception of auditory events has been termed the auditory horizon. Gardner (1969b) showed that the characteristics of the input (i.e., whispered vs. shouted speech) affect the distance estimate, even though amplitude loss with distance is the same for both inputs. The results indicated that the estimated distance of shouted speech tends to be greater than that for whispered speech. Coleman (1962) found that distance judgements were unrelated to actual distance on the first trial of a repeated-measures experiment. However, with repeated exposure, accuracy increased. This result indicates the importance of experience in judging distance.

The frequency spectra of a complex sound field has also been implicated in providing usable cues to distance. Changes in the frequency spectra occur as a function of the distance of the source from the observer. At short distances (< 4 feet) the sound field is relatively spherical in shape. Bekesy (1960) reported that a sound source appears to approach the observer if the low frequency components of the source are increased, relative to the high frequency information.

At farther distances the wave front becomes planar in shape (i.e., far-field conditions obtain). In the far-field, high frequencies are attenuated more rapidly than are lower frequencies as source distance increases. The difference in the rate of attenuation of high verus low frequencies depends on humidity, temperature, terrain features, and inhomogeneities in the atmosphere. These atmospheric effects are in addition to the $1/R^2$ loss.

The psychophysical data suggest, at least in monaural situations, that frequency composition of a sound serves as a relative cue for judgements of distance. Coleman (1968) showed that observers were accurate at localizing the distance of a source, located 8 to 24 feet away, based upon spectral cues. Loundsbury and Butler's (1979) investigation of spectral cues indicated that some observers were quite accurate at locating distances. However, 19% of their observers consistently inverted their estimates of distance.

Binaural cues to distance have also been investigated. The potential cues include interaural time differences (ITD's), interaural amplitude differences (IAD's), and interaural spectral differences (ISD's). Holt and Thurlow (1969) demonstrated that when interaural cues are available, observers can accurately rank-order the distances of the sources. When the source direction was to the side of the head, observers could accurately judge distance; but when the source was in front of the head, they could not.

Levy and Butler (1978) directly manipulated ITD, IAD and ISD cues. They recorded broad-band, low-pass, and high-pass noise bursts in stereo from a live model's head with the source at a distance of 5 feet. Three separate recordings were made which contained (1) ITD's, IAD's and ISD's, (2) only

ITD's and ISD's, or (3) only ITD's. The processed stimuli were then presented dichotically to a listener who judged their apparent distance. Judged distance was related to the frequency composition of the stimuli. Elimination of lower frequencies (and thus the fine-structure ITD cue) resulted in serious underestimation of the source distance. Distance estimates did not change significantly when IAD's and ISD's were eliminated. The results suggest that ITD's are the major binaural cue for distance perception. In a second experiment, they separated ongoing (fine structure) ITD's from onset (envelope) ITD's. Elimination of onset ITD's resulted in significantly greater underestimation of the source distance. Apparently both fine structure and envelope ITD's play a role in distance perception.

The effects of head movements, to create changes in IAD's, has been suggested as a binaural cue for distance. The approach employed has been to mathematically model the information required to derive distance. Hirsh (1968) attempted to model auditory depth perception based solely on interaural time and intensity differences. Unfortunately, without the repeated sampling available with head movements, Hirsh's model resulted in correlated equations, whereby IAD and ITD could not be independently estimated (Molino, 1973). Lambert (1974) has shown that listeners could judge distance based on the rate of change of the IAD as the head is rotated in the horizontal plane. However, no psychophysical evidence is yet available to support the usefulness of this cue or other head-movement related cues in judging distance.

2.5 Effects of Head Movement and Vision on Auditory Localization

Four major explanations have been advanced to account for the finding that head movements and the availability of visual inputs make auditory localization judgments more accurate. The four explanations should be regarded as different aspects of the same general problem, rather than as mutually exclusive hypotheses. They include: (1) the cue-modulation position (Wallach, 1940), (2) the central auditory acuity hypothesis (Pollack and Rose, 1967), (3) the visual frame-of-reference theory (Warren, 1970; Platt and Warren, 1972), and (4) the motor feedback/spatial memory position (Jones, 1975; Jones and Kabanoff, 1975).

The first explanation, advanced by Wallach (1940), is that moving the head, and hence the ears, modulates the binaural acoustic cues, i.e., interaural time and amplitude differences (ITD and IAD). Wallach noted that the change in these binaural cues with head movement depends on the initial direction of the sound source relative to the head. As the head is rotated counter-clockwise in the horizontal plane, the ITD and IAD increase for sources on the right-front and left-rear sides of the head, but decrease for sources in the other two quadrants. Pivoting the head from side to side produces corresponding effects in the vertical plane. Wallach suggested that such movements are especially helpful in resolving sound sources in front from those in back and those above from those below.

Lambert (1974) has developed a mathematical model which describes how the ITD cue changes as a function of head rotation for an idealized observer. The head is modeled as a sphere, with two holes for ears located at $\pm$ 90° azimuth. Assuming that the sound is of low enough frequency so that phase ambiguities do not occur, Lambert's model shows that the observer can uniquely determine the source azimuth based solely on the modulation of the ITD with head rotation.

Other acoustic cues also change as the head is rotated or pivoted. The change in IAD as a function of the azimuth of the source relative to the head was described in Section 2.2.2. The role of IAD in vertical localization is covered in Section 2.3 As noted in those sections, the IAD is sizeable only at medium and high frequencies. Studies of auditory lateralization using dichotic presentation have amply demonstrated that variation of the IAD causes the perceived location of the sound to shift laterally (cf., Mills, 1972; Durlach and Colburn, 1978).

Head rotation and pivoting should also modulate monaural (chiefly pinna-produced) cues. Under normal listening conditions, it is natural to reorient one's head toward a newly heard sound source. Freedman and Fisher (1968) have suggested that monaural cues help direct the initial orienting response (head movement) to a sound, but do not serve as important cues during head movement. They measured localization accuracy in the horizontal plane with listeners using their own pinnae, no (i.e., occluded) pinnae, or artifical pinnae, with or without head movement. With no head movement, their own or artifical pinnae increased localization accuracy considerably relative to the no-pinnae condition. However, when head movement was allowed, there was no significant difference between the pinnae and no-pinnae conditions. This result suggests that the modulation of binaural cues during head movement is so powerful that it swamps out any effect of monaural cue modulation. Unfortunately, Freedman and Fisher did not measure the time required to localize. If their hypothesis that pinnae cues help direct the initial orienting response is correct, then localization time should be shorter when pinnae cues are available.

A second hypothesis has been advanced by Pollack and Rose (1967). They found that head movement facilitated auditory localization only when (1) the sound source is initially located toward the side of the head, and (2) the duration of the stimulus exceeds the time required to reorient the head toward the source. They draw an analogy between foveal visual acuity and auditory acuity in the region around 0° azimuth, and suggest that head movement allows the listener to take advantage of the higher-acuity region.

A third explanation has been called the visual frame-of-reference hypothesis (Warren, 1970; Platt & Warren, 1972; Shelton & Searle, 1980). In the original formulation of this position, Warren (1970) suggested that, for adults, vision is the primary means of organizing sensory space, and that the necessary condition for facilitation of auditory localization is structured visual input. In a later article, Platt and Warren (1972) modified the original position to stress the interaction of eye movements and visual (i.e., retinal) information. They reported that eye movements in a lighted, textured environment facilitate localization relative to a condition in which the eyes are fixated in a lighted environment. However, eye movements in a dark environment did not produce a facilitation relative to the same control condition. Similar results have been reported by Mastroianni (1983b).

The facilitative effect of vision in the Platt and Warren (1972) study should be distinguished from the intersensory bias or "visual-capture" phenomenon. The latter phenomenon is the tendency for a sighted listener to perceive the origin of a sound to be a plausible visual object. In the Platt and Warren (1972) study, the visual background was a burlap curtain, i.e., there were no visual objects which could have been interpreted as the sound source. Pick, Warren, and Hay (1969) reported an experiment in which subjects pointed at (1) the <u>heard</u> position of a sound with the loudspeaker viewed through a prism which displaced the visual image, or (2) the <u>seen</u> position of the loudspeaker emitting sounds. The heard position of the sound was strongly biased in the direction of the displaced visual image, but the seen position was not biased by the sound.

A fourth explanation for the facilitative effect of vision on auditory localization has been advanced by Jones (1975) and Jones and Kabanoff (1975). They contend that motor activity alone provides the spatial framework, and that all sensory stimulation maps onto this spatial framework. These investigators suggest that eye movements in the direction of the auditory stimulus facilitate localization by "stabilizing" or updating spatial memory. Jones and Kabanoff explain Warren's (1970) and Platt and Warren's (1972) findings that eye movements facilitate auditory localization only in a lighted, textured environment by noting that eye movements are more accurate in the light. They propose that the role of retinal information is to reduce eye drift rather than to provide the primary spatial point of reference. Jones and Kabanoff (1975) report an experiment in which eye movements in the direction of an auditory stimulus facilitated localization, whereas eye movements in the opposite direction produced less accurate localization than a control condition in which the eyes were fixated. In that experiment, the measure of localization accuracy was the observer's percentage of correct judgements as to whether the auditory stimulus was 3° to the right or 3° to the left of the median plane of his or her body. The speakers were hidden from view in order to rule out visual capture effects. The finding that eye movements in the direction opposite the sound degraded localization is contrary to the visual-frame-of-reference position. However, this finding does not rule out the possibility that visual (retinal) information may play some role in conjunction with eye movements to facilitate auditory localization.

Shelton, Rodgers, and Searle (1982) note that the visual-frame-of-reference and the spatial memory positions are not mutually exclusive. Even though the eyes are constantly moving, one perceives stationary objects in the environment as remaining stationary. The fact that the image of an object moves across the retina does not necessarily lead to the perception that the object is moving. The visual system must somehow integrate retinal information and extra-retinal eye-position information (EEPI) in order to determine whether the object or the eye is moving (Matin, Stevens and Picoult, 1983). Given that retinal information and EEPI are used in visual localization, it seems plausible that they both also play a role in auditory localization.

It is also clear that the modulation of acoustic cues with head motion plays a role in auditory localization independently of vision. Thurlow and Runge (1967) found that head movements facilitated auditory localization even when subjects where blindfolded. Wallach (1940) reports three conditions in which a blindfolded listener was passively rotated in a chair and a sound source was also rotated about the same axis, starting from a point directly in front of the listener. Depending on the rate of rotation of the sound relative to that of the listener, several different auditory illusions were created. When the sound rotated at the same rate as the listener, it was perceived as emanating from above the listener's head. When the sound was rotated at twice the rate of the listener, it was perceived as coming from behind, and when it was rotated at 1.5 times the listener's rate, it appeared to be at an elevation of 60° and behind.

Logically, it seems necessary that the brain must integrate the acoustic cues produced by head motion with proprioceptive, vestibular, and/or visual cues. A sound may move relative to the ears either because the head is rotating and the sound source is stationary, or because the sound is moving relative to the body and the head is stationary relative to the body. In order for a listener to distinguish between these two cases, the changing acoustic cues must be integrated with proprioceptive feedback from the neck muscles, and/or vestibular cues and/or visual feedback.

Figure 18 shows a model of auditory or auditory <u>and</u> visual localization which provides an integrative framework for the evidence offered in support of all four of the positions discussed above. The model shows how visual, vestibular, proprioceptive, and auditory information might interact in an auditory localization task. The model is made up of three feedback control loops, each of which consists of an error detection process and a process being controlled. The error detector combines feedback from the controlled process with other reference information to generate an error signal. The error signal guides the controlled process. Control loop C represents the process of visually fixating or tracking an object in space. The error signal from this control loop serves as one of several kinds of reference inputs to control loop B. Loop B represents the process of orienting the head toward

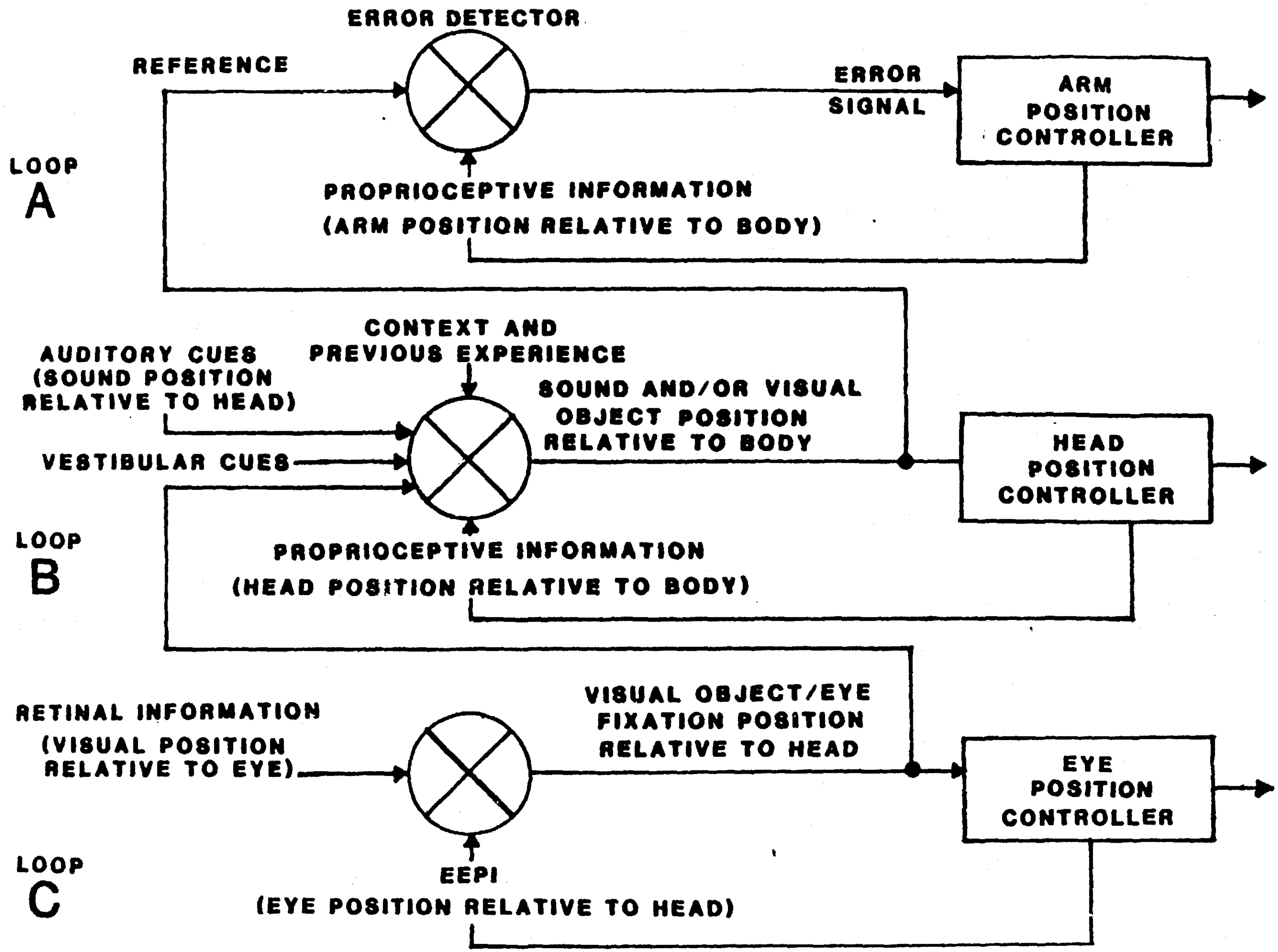


Figure 18. Feedback control model of auditory localization.

(or tracking) a visual or acoustic event in space. Depending on the situation, auditory, vestibular, visual, and proprioceptive cues may all play a role in determining the error signal in this control loop. Contextual cues and prior knowledge may also play a role. For example, Welsh and Warren (1980) argue that intersensory bias occurs only when the listener assumes that the auditory and visual stimuli are representative of the same object.

Control loop A in Figure 18 represents the process of pointing the arm to localize a sound source. The reference input for this loop is the error signal from loop B, i.e., the sound source direction relative to the body.

Control loop B requires at least one reference input and proprioceptive feedback from the neck muscles. The sole reference input must, of course, be appropriate to the modality of the stimulus being localized. Thus a blindfolded, static listener can localize a sound even though vestibular and visual cues are not available.

When more than one reference input is available, it seems reasonable to expect that the accuracy of the error signal generated by the error detector in loop B may be enhanced. If a secondary reference input is more precise than the original, as vision is relative to audition, then auditory localization accuracy should be enhanced, as the studies cited previously have demonstrated.

The Jones and Kabanoff (1975) finding that eye movements in the direction opposite an auditory stimulus degrade auditory localization may be explained as follows. Eye movements in the opposite direction should generate reference inputs to control loop B which are irrelevant to the task at hand. If they cannot be ignored, then they might well decrease the accuracy of the error signal generated in that loop. In the Jones and Kabanoff study, eye movements were cued away from, versus toward, the sound source randomly from trial to trial. Therefore, the subject had no way of knowing a priori whether the eye movements were consistent or inconsistent with the location of the auditory image. Thus, it seems that irrelevant visual information can disrupt the process of judging the sound direction relative to the body, which is represented by control loop B.

The feedback-control-loop model suggests that the listener should be able to ignore irrelevant (or even inconsistent) visual information during auditory localization if he or she knows that it is irrelevant. Easton (1983) reports a study in which inconsistent visual information was successfully ignored in auditory localization. Subjects located a small loudspeaker using either the visual, auditory, or both modalities. On dual-modality trials, subjects viewed the speaker through a prism which displaced the visual image. On one-half of the trials the subject was to point to the <u>seen</u> position of the speaker; on the other trials he/she was to point to the <u>heard</u> position. In different conditions, the subject's head was immobilized or free to move, and the subject was either unaware of the visual distortion or was informed. When the head was immobile, informing subjects of the visual distortion greatly reduced visual bias of the auditory localization response. When subjects were allowed free head movement, informing them of the visual distortion completely eliminated the visual bias in localizing the heard position of the speaker.

The feedback-control-loop model of auditory localization was suggested in part by the work of Matin and his associates (cf., Matin, 1982a, 1982b; Matin, Picoult, Stevens, Edwards, Young, and MacArthur, 1982; Matin, Stevens, and Picoult, 1983). Their work shows that the feedback from eye movements (i.e., EEPI) necessarily plays a role in visual localization, and can also play a strong part in auditory localization. In one experiment, Matin et al. (1982, 1983) partially paralyzed observers by administering curare. The curare had the effect of subjectively requiring increased effort to move the eyes but did not affect auditory localization accuracy. With the head immobilized, an observer was told to fixate a visual target at various azimuths ($\pm$ 10°, $\pm$ 15°) in the horizontal plane, at eye level. The observer then judged when a sound, which could be switched among any one of 25 loudspeakers spanning $\pm$ 30° azimuth, was co-located with the fixated light. Paralyzed observers consistently placed the sound at a greater absolute azimuth from the median plane than the fixated light. Unparalyzed observers were able to match the light and sound positions accurately. This result suggests that EEPI is distorted in paralyzed observers; specifically, that greater effort is required to move and hold the eyes at a given deviation from the median plane of the head. It appears that the observer interprets the greater effort to mean that his eyes are at a greater angle from the median plane than they actually are. This causes the observer to place the sound at that perceived greater angle, which exceeds the actual angle of the light. Matin et al. conclude that EEPI is used to judge the position of the eyes relative to the head, and also serves as a spatial frame of reference for auditory localization.

The feedback-control-loop model is a useful device for integrating the many findings concerning head and eye movement effects on auditory localization. It should be noted that the model is intended to account only for the effects of other sensory inputs on auditory localization. It may not apply to purely visual localization. The model may provide a useful framework for further research. Some implications of the model for further research are discussed in Section 2.9.

2.6 Perception of Auditory Motion and Volume

The preceding section on head and eye movements makes obvious the importance of evaluating auditory localization in dynamic situations. Another facet of localization in dynamic situations involves locating sources which are themselves in motion. Auditory motion may be defined relative to both real or illusory motion, and to the general location of the stimulus (e.g., intracranial movement or external source movement).

The investigation of illusory auditory motion has paralleled the study of apparent movement in the visual system. Perrott (1982) reviewed three types of illusory motion: the auditory autokinetic effect, motion after-effects, and induced motion. The report of auditory movement, or jitter, of a stationary source has been termed auditory autokinesis. Perrott, Mason and Forbes (1973), in their review of the auditory autokinesis, identified the effects of signal bandwidth on the illusion. Increases in bandwidth result in decreased incidence of perceived motion and decreased duration of motion if it does occur. Greater bandwidth also increases the latency to the first perception of movement. The auditory motion aftereffect and its visual counterpart, the waterfall illusion, create the perception of movement through the termination of a moving auditory background. The perception of movement occurs in the opposite direction of the background. Perrott (1982) reported that only minimal target displacement (a few degrees) has been observed. Induced motion creates the perception of movement through the movement of nearby sources. The effects of induced motion effects are reportedly small (Perrott, 1982).

Perrott and Musicant (1977, 1981) studied listener's sensitivity to real motion of a sound source in the horizontal plane. Perrot and Musicant (1977) measured the minimum arc through which a sound must travel in order to be discriminable from a stationary source at roughly the same azimuth. The duration of a 500 Hz tone emitted from a loudspeaker rotating about the listener was adjusted until the listener correctly discriminated the moving source from a stationary source on 75% of the trials. The minimum arc required for discrimination was called the minimum audible movement angle (MAMA). The MAMA is, of course, confounded with signal duration. The MAMA increased linearly with angular velocity: from 8.3° at 90°/sec to 21.2° at 360°/sec.

In a second study, Perrott and Musicant measured a dynamic counterpart of the minimum audible angle (MAA) developed by Mills (1958). Observers adjusted the onset of a 500 Hz tone emitted by a loudspeaker rotating about them in the horizontal plane until it corresponded with the position of a light (at 0° azimuth). The standard deviation of the position at which the sound was turned on (the dynamic MAA) was about 1° for sound velocities of 45, 60, and 120°/sec. The dynamic MAA was about 2.5° at a source velocity of 240°/sec. However, the mean position at which the sound was turned on was biased in the direction of rotation of the sound (about 5° at 45°/sec to 12° at 240°/sec). Perrott and Musicant note that the dynamic MAA corresponds in magnitude to the static MAA obtained by Mills and that the MAMA (Perrott and Musicant, 1977) is much larger. The difference, they suggest, indicates that the MAMA measures a fundamentally different capability of the auditory system than does the MAA.

Perrott (1982) examined observer's ability to judge the velocity of a sound source in a darkened chamber. For comparison purposes, observers also judged the velocity of a silent source visually, with the chamber illuminated. Both the velocity and duration of band limited (100 - 1000 Hz) noise were varied. Observers were quite consistent in their estimates of velocity. Judgements of auditory velocity closely paralleled those of visual velocity. Both followed a simple power law with a slope slightly less than 1.0; although for both modalities absolute velocity was overestimated. In addition, Perrott found that the duration of a moving stimulus has to exceed 100 msec in order for the observers to detect motion. It was concluded that the auditory system is as capable of detecting velocity as is the visual system.

Illusory auditory motion also occurs under some conditions in lateralization tasks. The intracranial position of an auditory stimulus is dependent upon the interstimulus onset interval (ISOI). That is, the interval between the onsets of dichotically presented stimuli determines whether two separate auditory events are heard, or whether a single auditory event (phantom source) is heard. The fused phantom sound can either be perceived as stationary or as moving from ear to ear. Briggs and Perrott (1972) varied the ISOI in a dichotic listening paradigm and asked subjects to report the position(s) of the stimuli. For 100 msec signal duration, ISOIs less than 15 msec tended to produce a single fixed source image. For ISOIs between 15 and 75 msecs, listeners perceived a single source which moved continuously from the lead to the lag ear. ISOIs between 75 and 110 msec produced a fused image whose movement could best be described as broken; and at ISOIs greater than 110 msec the stimuli were heard as two successive sounds located at each ear. Perrott and Baars (1974) varied the signal duration and found that as duration increased, the threshold for each of the above effects shifted towards longer ISOI values. That is, the perception of a fused, continuously moving image shifted in ISOI from a mean of 19 msec for a 10 msec signal to 62 msec for a 300 msec signal. These results indicate that apparent movement is largely a function of the ISOI and signal duration.

Closely related to auditory movement in lateralization tasks is the perception of the spatial extent of an auditory image. Spatial extent can be defined as the amount of intracranial area, or volume, that an image is perceived to occupy. Spatial extent is a direct function of the degree of coherence, or correlation, between two input signals (Blauert, 1982; 1983). Fully coherent signals ($r = 1.0$) are perceived as a fused auditory image with limited spatial extent. As coherence decreases, from 1.0 to approximately .4, the area of the auditory image tends to increase. Signals below .4 coherence are perceived as two separate events, each with limited extent. Blauert (1983) hypothesized that the auditory system must perform some sort of cross-correlation analysis, where the correlation between the signals provides information about the location and extent of the source.

2.7 Effects of Noise on Auditory Localization

In natural settings most judgements of sound location are made in a background noise. Research has been conducted to assess the noise characteristics that interfere with or mask the location of the sound source, and the changes in localization accuracy that occur in the presence of noise. Noise effects, in lateralization and masking paradigms, have provided some additional evidence on noise effects in localization. These effects will be reviewed as they pertain to localization in the presence of noise.

The studies on auditory localization have assessed the impact of noise on localization accuracy, the minimal audible angle, and mask frequency effects. Jacobsen (1976, cited in Blauert, 1983) investigated the effects of noise on localization accuracy, and found that localization was as accurate in the presence of noise (automotive traffic at 30 sones) as in quiet, if the sound source was from 10 to 15 db above masking threshold. An early study by King and Laird (1930) identified the effects of noise on the minimum audible angle for click stimuli in the frontal part of the horizontal plane. The difference threshold increased from 1.8 degrees in quiet to 4 degrees with 45 db noise present.

Moter (1964, cited in Durlach and Colburn, 1978) varied the angular separation between the target and masker, and measured the detection threshold. The detection threshold decreased as the angle between the target and masker increased. Kock (1950) investigated the threshold of detectability for speech in a white noise background as a function of interaural time differences. When the signal and noise arrived at the ears with the same interaural time delay, the threshold for the speech signal was between 8 to 16 db greater than when the interaural delays differed. These results were interpreted as an "unmasking" of the signal as the interaural time difference between the signal and the noise mask increased. Explanatory models of this binaural "unmasking" phenomenon have been developed by Jeffress (1972) and Durlach (1972).

Canevet, Germain and Scharf (1980) investigated the effect of signal masking as both the signal and mask frequencies were varied. Critical frequency bands of maximal masking were identified with regards to the signal frequency. In general, as the frequency band of the mask overlaps with the signal frequency, localization accuracy decreases. The effective critical bandwidth of a mask tended to increase as the signal frequency increased, and the critical bands were not symmetrical around the signal center frequency. Canevet, Germain and Scharf (1979) and Scharf, Canevet, Buus, and Marchioni (1982) presented a mask which preceded signal onset, within the bounds of the precedence effect. The masking stimulus was close in frequency to the signal. They found that the localization threshold was 10 to 20 db greater than the signal detection threshold. As the frequency difference between the signal and mask increased, the detection threshold decreased more rapidly than did the localization threshold.

The effects of noise on the lateralization of a signal have been more extensively studied (Gaskell and Henning, 1979; Ito, Thompson, and Colburn, 1979; McFadden, 1969; Robinson and Egan, 1974; Wilbanks, 1983). In general, the presence of noise tends to reduce lateraiization accuracy when the interaural time envelopes of the signal and noise overlap. These results have been interpreted in terms of phase cancellation and reinforcement. When the signal and noise bands overlap in frequency, phase cancellation and reinforcement increase; therefore the noise tends to interfere with detection and lateralization more.

The ability to localize sound in the presence of background noise is often observed in daily life. The aptly named "cocktail party effect," (Batteau, 1968; Durlach and Colburn, 1978) adequately characterizes the ability to localize and attend to a source amongst a din of background noise. The evidence presented above has identified the effects of background noise on localization, lateralization, and detection as a function of the angular separation of the signal and noise sources, signal and noise frequency, and interaural time differences. However, further research is needed to identify signal characteristics which lead to minimal disruption of auditory localisation in noise.

2.8 Cockpit Applications of Simulated Auditory Localisation

The foregoing review of research suggests that the auditory modality could be used to convey accurate spatial information. The human observer's ability to localise auditory signals is completely neglected and unused in modern aircraft. Several previous attempts to develop auditory displays of spatial information for aircraft never found any application (cf., Forbes, 1946; Mudd, 1965; Mudd and McCormick, 1960). However, in the twenty years since Mudd's study, our knowledge of auditory localisation has improved tremendously. The availability of this new knowledge presents an opportunity to relieve the visual modality, which is currently over-taxed in cockpit settings.

This section addresses three areas in which spatial information conveyed by way of audition could potentially be used to improve aircrew performance. The first of these, directional cueing in head-coupled control/display systems, is the main focus of the present project. A second area, of which the first is a special case, is the enhancement of the pilot's situational awareness. A third area is the enhancement of audio communications and audio warnings through exploitation of the human's ability to attend selectively to spatially separated messages.

2.8.1 Directional Cueing

The sensitivity and range of advanced electromagnetic (EM) sensors aboard modern military aircraft far exceed the capabilities of the pilot's unaided senses. Modern sensors such as radar, infrared (IR) and electro-optical (EO) systems can each provide a separate, amplified view of the outside world beyond the range of the naked eye. Unfortunately, the pilot is often over-loaded with visual information from a multitude of cockpit displays, and therefore cannot effectively use all the information available from these sensors. Furthermore, because space available on the cockpit display front panel is severely limited, the displays associated with each sensor must be made so small that only a part of the information available can be effectively conveyed to the pilot.

These problems have been partially alleviated by integrating some of the systems which provide information and by presenting integrated information on head-up displays (HUD's), and multifunction displays (MFD's). The effectiveness of these systems depends in large part on how effectively the information is integrated. At the lowest level, two kinds of information, one highly symbolic and the other concrete, are simply superimposed, as when flight control symbols are overlayed on the outside visual scene using a HUD.

At a more advanced level, a concrete picture of the outside world representing threat envelopes, the desired flight path, etc., can be generated on an MFD and overlayed on the outside visual scene by projecting the picture onto a HUD. One of the most significant aspects of this type of system is its ability to represent spatial information from various avionics systems in the most concrete form possible, as (virtual) objects in visual space. Unfortunately, the MFD-HUD picture is confined to the angular region around the longitudinal axis (boresight) of the aircraft. This means, for example, that a reticule on the HUD can be used to aim a weapon only when the target is boresighted with the aircraft.

The Head-Coupled Control/Display technology now under development at AAMRL/HEA takes the next logical step beyond the stationary HUD. Information from avionics sensors is represented as concrete objects in a panoramic visual picture. This is accomplished by presenting a virtual image on the pilot's helmet-mounted visor. The image projected on the visor changes as the pilot rotates his or her head, just as the real world scene normally changes as we turn our heads. This system, called the helmet-mounted display (HMD), presents an omnidirectional view of the outside world enhanced by information from avionics sensors. The pilot can also use his or her head position to designate a target by aligning a reticule, which remains stationary on the visor, over the desired target. Magnetic sensors on the pilot's helmet provide continuous measurement of the pilot's head orientation, and therefore of the pilot's line of sight when the eyes are positioned straight and level. This sytem is called the helmet-mounted sight (HMS).

To date, the Head-Coupled Control/Display technology program has focused on the visual modality. Although vision is our primary means of obtaining spatial information about the world around us, audition also plays a significant role. In the natural environment, sounds are often a cue as to the direction of important objects or sources of information. Upon hearing a novel sound, a human generally rotates his or her head and eyes toward the source and acquires it visually. Aural information is a natural cue for where to look in our visual environment.

Simulated auditory localisation (SAL) could be used to cue the pilot where to look in the virtual visual world provided by the Head-Coupled Control/Display system. The system currently conveys directional information by presenting visual symbols, such as arrows, on the HMD. Directional information could instead be coded as appropriate aural signals to the ears. The acoustic cues critical to auditory localization can now be simulated and presented dichotically, via earphones, with sufficient fidelity to achieve localization performance comparable to that with the unaided ear (Batteau, 1965; Weinrich, 1982). Batteau generated acoustic localization cues by recording from miniature microphones placed in a physical model of the head and pinnae. Obviously, an electronic simulation would be necessary to make SAL practical for cockpit applications.

Fortunately, the knowledge is now available to make at least a "brute-force" electronic simulation feasible. The brute-force method would involve storing a large number of transfer functions. Each function would describe the relationship of the free-field sound (if the observer's head were not present) to the proximal stimulus at the entrance to the ear canal for each apparent position of the sound source to be simulated. Obviously, a large number of positions would be involved, and an algorithm for transitioning between positions would be needed in order to allow for head movement. The brute-force method could be implemented by using waveform digitization methods, like those used in speech synthesizers. Obviously, a rule-based simulation would be more elegant. The review of research on auditory localization presented in Sections 2.2 to 2.7 suggests that a rule-based system will be possible in the near future.

There are at least three ways in which SAL could facilitate a pilot's visual target acquisition and overall performance. First, using SAL instead of visual indicators should reduce the workload in the often over-used visual modality. Second, the visual orienting response to a localized auditory stimulus is natural and highly automatic, which should make it faster and more compatible with competing tasks, such as flight control (cf., Posner, 1978; Shiffrin & Schneider, 1977). Third, recent research on auditory localization indicates that auditory localization can be highly accurate if the listener is allowed to reorient his or her head and eyes toward the source (see Section 2.5). It may be possible to reorient the head and eyes more accurately with SAL than with symbolic visual indicators, thereby reducing the number of eye movements required to acquire a target. These considerations suggested that SAL should be especially valuable in facilitating rapid acquisition of visual targets.

2.8.1.1 Performance Implications of SAL Directional Cues

Further research is needed to investigate issues related to the feasibility and potential benefit of using SAL as a directional cue in the cockpit. As implied above, these issues include the effectiveness of SAL relative to other methods of providing directional information for a visual acquisition task, and the relative time-sharing efficiency of SAL with competing flight tasks. For example, SAL should be compared to conventional cockpit visual displays (CVD's), HUD's, HMD's, and synthesized speech displays (without acoustic location cues). Both speech and non-speech SAL signals should be compared, and all of the directional cueing methods should be evaluated at various levels of workload imposed by various types of concurrent tasks.

Wickens, Sandry & Vidulich (1983) have proposed an extension of the stimulus-response (S-R) compatibility principle to include a central processing (C) component. The S-C-R principle postulates that certain combinations of input modality, output modality, and type of central processing (i.e., verbal, spatial) result in better task performance than do others. According to the principle, tasks which require spatial processing and a manual response should be performed better when the input modality is visual than when it is auditory. The principle predicts that a CVD, HUD, or HMD should produce faster visual acquisition than SAL. However, orientation of the eyes in response to auditory localization cues is a highly natural, perhaps "automatic" process (cf., Posner, 1978). SAL cues may therefore produce faster visual acquisition than do the CVD's, HUD's or HMD's. Such a finding would represent an important exception to the S-C-R compatibility principle.

Not only should SAL produce better performance than symbolic visual displays (i.e., CVD's, HUD's, and HMD's) in single-task situations, but it should also be superior in multitask settings. Two lines of reasoning suggest that a SAL-visual acquisition task should be time-shared more efficiently with a concurrent visual-spatial-manual task than are visual acquisition tasks based on symbolic visual displays.

First, the multiple-resource theory of task interference (c.f., Navon & Gopher, 1979; Norman & Bobrow, 1968; Wickens, 1980) suggests that the time-sharing efficiency of concurrent tasks decreases as the amount of overlap of input modalities, type of central processing, and output modalities increases. Therefore, increasing concurrent visual workload during a visual acquisition task should affect visual acquisition time more when directional information is provided by a CVD, HUD, or HMD than when it is provided by SAL cues, since SAL involves auditory input.

Second, auditory localization is also probably more natural or automatic than is localization based on a symbolic visual display. This also leads one to expect that visual acquisition time should be more affected by competing visual workload in the CVD, HUD, and HMD conditions than in the SAL condition. When directional information is provided by synthesized speech, increased workload may affect visual acquisition time to an intermediate extent, since localization on the basis of speech is not automatic, but on the other hand, the input modality differs from that of the competing task.

With a concurrent auditory-verbal-speech task, such as cockpit communications, multiple-resource theory predicts that the CVD, HUD, and HMD should produce faster visual acquisition than would SAL, since SAL and the concurrent task share the same input modality. The high automaticity of auditory localization suggests, however, that it may be little affected by competing workload (cf., Posner, 1978; Shiffrin & Schneider, 1977). Contrary to multiple-resource theory, this leads to the expectation that visual acquisition time with SAL directional cues may be less affected by competing workload than the HUD, HMD, and CVD conditions. Such a result would be an exception to the multiple-resource theory of task interference, and would therefore have implications for task and human-machine interface design.

2.8.1.2 Other Issues Affecting the Usefulness of SAL Directional Cues

Other issues related to the viability and effectiveness of using SAL to provide directional cues for visual acquisition in the cockpit include: (1) the accuracy of SAL in noise, (2) the rapidity with which listeners can adapt to SAL cues, and (3) the effectiveness in terms of localization speed and accuracy of various types of free-field SAL stimuli. This last issue was addressed in the experiments reported in Section 5.0 of this report. Findings discussed previously in Section 2.2.3 suggest that listeners adapt to SAL cues within minutes. The findings reported in Section 2.7 indicate that the impact of noise on SAL accuracy will depend on the bandwidth and frequency composition of the stimulus. The larger the number of frequency components in the free-field SAL stimulus, or the greater the bandwidth, the more resistant it should be to masking in varying noise conditions. Research is needed to quantify exactly how much typical aircraft noise degrades SAL performance and to identify the properties of SAL stimuli which best retain their usefulness in noise.

2.8.2 Enhancement of Situational Awareness

The use of SAL to cue the pilot where to look in the visual environment is a special case of using SAL to enhance the pilot's situational awareness. In addition to redirecting the pilot's visual attention, however, enhancing situational awareness also implies keeping the pilot informed while making fewer demands on visual attention. For example, auditory cues could provide information about terrain clearance and obstacles in low-level flight. Auditory cues could alert the pilot when a dangerous condition develops, and therefore reduce the amount of visual fixation time and visual processing the pilot spends monitoring the situation. For example, the distance to an approaching obstacle could be cued by the intensity and frequency composition of a sound. Change in interaural amplitude differences with head movement might also provide a useful cue for distance (see Section 2.4). Different obstacles would, of course, be perceptually separated by providing different acoustic cues for azimuth and elevation. The findings reviewed in Section 2.3 suggest that frequency could be used as a cue to the elevation of an obstacle. The relative speed of approach toward an obstacle could be cued by the rate of change of intensity and rate of change of frequency composition. Approach speed might also be cued by a general shift in frequency, simulating the Doppler effect.

At a more advanced level, simulated and enhanced echo-location cues could be used to help the pilot judge distances to obstacles in nap-of-the-earth flight, supplementing visual cues. It might also be possible to use SAL to help maintain the pilot's spatial orientation when flying in low visibility conditions (i.e., as an anti-vertigo cue). Appropriate use of auditory cues for functions such as these might greatly enhance the pilot's situational awareness with very little expenditure of mental processing resources.

The major reason for using audition to enhance situational awareness would be to facilitate quick recognition and correction of dangerous flight conditions when the pilot's attention is directed elsewhere. The appropriate test of such cues is, therefore, in high workload conditions.

2.8.3 Enhancement of Communications and Audio Warnings

A well known aspect of audition is binaural "unmasking" or the binaural masking level difference (MLD). The sensitivity of the auditory system for a signal in a noise background is much higher for dichotic listening conditions than in diotic conditions. For example, presenting identical 500 Hz signal and noise stimuli to both ears, but reversing the phase of the signal (but not the noise) in one ear, lowers the signal threshold by 11 db (Jeffress, 1972). The MLD is smaller at high frequencies, e.g., about 3 db at 2 kHz and at 5 kHz. Durlach (1964) proposed that the basis for binaural unmasking is interaural phase or time differences at low frequency (< 1500 Hz) and interaural amplitude differences at high frequencies.

The well known "cocktail party" effect is, of course, an example of binaural unmasking. An audiometrically normal listener can follow one voice amidst a background of other conversations and noise. However, if the same conversations and noise are recorded monaurally and played back, it is very difficult to follow any one voice. The apparent location of the source helps the listener discriminate among the voices. It is clear from lateralization research that binaural differences contribute to this discrimination ability (cf., Mills, 1972). It is apparently not known whether, or to what extent, monaural localization cues also contribute to the ability to discriminate a spatially distinct signal from noise.

Aircraft communications and audio warnings are generally binaurally unmasked from ambient noise by reversing the signal phase at one earphone. However, this does not separate the signal from noise or other concurrent signals (such as communications jamming) coming through the same audio system. The detectability and intelligibility of communications and audio warnings might be greatly enhanced if each signal were given a different apparent direction. This would be relatively easy to accomplish for threat warnings, since threat warning systems generally measure angle-of-arrival information. Aircraft audio warnings and messages coming over UHF as opposed

VHF radio could also be given different apparent directions. Research is needed to determine the benefit of associating cockpit signals with different apparent directions in conditions of noise and jamming.

2.9 Recommendations for Further Research

In addition to the applications-oriented research issues discussed in the last section, there are many basic research issues which should be addressed in order to facilitate the application of SAL in the cockpit. A few of these basic issues which seem most important are discussed in this section. An obvious need is for a specification of the necessary and sufficient cues for localization in both the horizontal and vertical planes. Although the binaural cues (ITD, IAD) can be specified in any given situation, much less is known about monaural cues. Isolation of the psychologically necessary and sufficient acoustic features for localization would greatly simplify the electronic generation of SAL cues.

A first task is to account for the discrepant results between listeners and test situations in the study of monaural cues. The results reviewed in Sections 2.2.3 and 2.3.2 show that narrow-band noise stimuli in the 4 to 12 kHz frequency range produce the illusion that the source is located in the median sagittal plane in some cases, and in the horizontal plane in other conditions. Moreover, some listeners map frequencies from 4 to 8 kHz and again from 8 to 12 kHz into the front quadrant of the horizontal or median plane, while for other listeners the spatial referents of frequencies from 4-12 kHz cover the entire hemisphere from front to back. These results suggest that: (1) context and prior knowledge play an important role in the interpretation of monaural cues, and/or (2) the narrow-band noise stimuli are not fully adequate cues for source direction (i.e., such stimuli do not include all the acoustic features produced by interaction of the free-field sound with the pinnae and torso). The experiments on monaural auditory illusions should be redone using noise stimuli with features that mimic the head-related transfer functions in the 4 to 12 kHz region. The features should be systematically tested to determine which are necessary and sufficient for monaural localization. Further transformation of the free-field sound spectrum by the pinnae should be eliminated by using headphone presentation. This research should also carefully control and/or manipulate the listener's expectations with regard to possible source locations.

A second related research question is whether monaural cues consist of narrow frequency bands which are amplified relative to the other parts of the spectrum, or "comb-filter" patterns like those produced by a multiple-delay-and-add system (see the discussion at the end of Section 2.3.2). Watkins (1978) produced the illusion of a moving sound source in the median plane by varying the delay of one component of a two-delay-and-add signal. Varying the delay produces a stimulus with a changing comb-filter spectrum. Stimuli with a fixed comb-filter pattern spectrum have apparently not been tested to determine whether they produce auditory illusions. Further research is needed to determine whether comb-filter spectral patterns produced by a multiple-delay-and-add-system with fixed delays produce the illusion of a source located at a fixed location in the median plane. Similar experiments should also be done in the horizontal plane. In the horizontal plane, the illusion produced by comb-filter spectral patterns in combination with appropriate ITD's and IAD's should be investigated. Watkins (1978) pointed out that the comb-filter patterns produced by a two-delay-and-add system with an appropriate range of delay values (based on Batteau's, 1967, analysis of pinna reflections) consist of a pattern of several peaks and notches in the audible frequency range (see Figure 17). In some cases a single amplitude peak (e.g. 12 kHz) is associated with two apparent source directions. Thus, if one simulated only a peak at 12 kHz, one might expect ambiguous localization, as the studies of auditory illusions with monaural stimuli indeed found.

A third area needing research is the investigation of acoustic cues which enable the listener to judge the distance of a source. There are at least two ways in which listeners might derive absolute information as to source distance through repeated sampling during head movements. Lambert (1974) has shown that listeners <u>could</u> judge source distance from the rate of change of the IAD as the head is rotated in the horizontal plane.

A second possible mechanism which listeners could use to judge distance is to compare the magnitude of the ITD to monaural pinna cues for source elevation and azimuth. For example, suppose a point source is located in the horizontal plane, at an azimuth of 90° (opposite the left ear). If the source were at infinity, the ITD would be:

$$\frac{3\pi}{2} \; \frac{r}{c} = 1.2 \text{ msec} ,$$

where r is the radius of the head and c is the speed of sound. As the source moves closer, monaural pinna cues remain relatively unchanged, but the ITD decreases. In the extreme case, where the source is located on the surface of the head, the ITD is:

$$\pi \frac{r}{c} = 800.\ \mu\text{sec}\ .$$

The rate of the change in ITD with source distance is largest when the source is at 90° azimuth and when the source is close to the head. It is not clear whether near-field acoustic effects would impair the usefulness of this cue, or whether the monaural cues are sufficiently accurate to allow the comparison required. It should be noted that this cue does not require head movement, but it does require a lateral source location and a broad-band stimulus. This and the head movement-related cue suggested by Lambert should be investigated experimentally.

The effects of head movement and vision on auditory localization is a complex area in which there are many issues that warrant further research. Two general issues will be raised here: (1) the nature of the cues which are modulated during head movement, and (2) the effect of other sensory inputs and proprioception on auditory localization.

Section 2.5 discusses three hypotheses which are variations on the cue-modulation position. The first of these hypotheses is Lambert's (1974) mathematical model, which shows that listeners <u>could</u> make absolute judgements of source azimuth in the horizontal plane based on the rate of change of ITD with the angle of head rotation, θ. Obviously, the modulation of other cues (e.g., IAD, pinnae cues) might also play a role in the facilitative effect of head movement. It would be relatively easy to test the cue modulation hypothesis by independently manipulating: (1) the characteristics of the stimulus, so as to control the availability of each cue (see Section 4.0), and (2) the presence or absence of cue modulations, by either restraining the listener's head or allowing free head movement.

Two additional hypotheses related to the cue modulation position could also be easily tested in the same experiment. Freedman and Fisher (1968) suggested that pinnae cues help direct the initial orienting response to a sound, but do not serve as important cues during head movement. As noted in Section 2.5, if their hypothesis is correct, then the time required to localize a sound source (i.e., response time in a speeded task) should be smaller when pinnae cues are available during head movement than when they are not. Pollack and Rose (1967) propose that the facilitative effect of head movement is due to the fact that the listener can take advantage of the high auditory acuity in the 0° azimuth region. This hypothesis could be tested by manipulating the amount and direction of head movement allowed during localization.

The feedback control model presented in Section 2.5 suggests three different areas of research related to the effects of head movement and vision on auditory localization. These areas include (1) research to evaluate the usefulness of the feedback control approach as an explanatory model, (2) research to evaluate the transfer functions relating inputs (reference signals) in the model to outputs, and (3) research which attempts to quantify individual differences in localization performance based on model parameters.

If the feedback control modeling approach is useful, then it should be possible to deduce and experimentally test previously untested hypotheses. An implication of the model, shown in Figure 18, is the existence of a purely auditory illusion analogous to the oculomotor - visual illusion reported by Matin, Picoult, Stevens, Edwards, Young, and MacArthur (1982) and Matin, Stevens, and Picoult (1983). Moving the head to localize a sound is somewhat analogous to moving the eyes in visual localization. A critical experiment would involve having an observer, partially paralyzed by the drug curare, rotate the head until he or she perceived a sound, located laterally from the median plane of the body, to be directly in front of the head. The observer would then be asked to locate a second sound in the median plane of the body. If the observer perceives that the head is turned farther than it really is, due to the increased effort required because of the curare, then the second sound should be positioned so as to deviate from the median plane in the direction opposite the head orientation. Such a result would replicate Matin et al.'s (1982, 1983) results with visual stimuli and provide evidence that proprioceptive cues associated with head movement also (like eye movement information) provide a spatial frame-of-reference for auditory localization. This experiment is essentially a test of the postulated control loop B of the feedback control model (see Figure 18).

If the general structure of the feedback control model can be verified, then it can be used as a tool for quantifying the effect of sensory inputs and proprioception on auditory localization. A control theory approach to intersensory effects in auditory localization would force investigators to examine relationships among sensory inputs and motor outputs which might otherwise be ignored. The approach would also result in the identification of closed-loop and intersensory transfer functions which would be useful in predicting auditory localization performance.

Studies of individual performances in the control-theory framework would make possible a more precise quantification of the basic skills underlying exceptional performance. Such information could be very useful in selection and training.

Two additional research areas which are important to SAL applications are auditory motion perception and the effects of noise on auditory localization performance. Further research is needed to determine the cues responsible for the perception of motion for auditory sources which are actually moving. Such information is, of course, crucial to simulating auditory motion. A salient question related to noise effects on localization is the extent to which typical aircraft background noise and jamming degrade the effectiveness of auditory directional cues. It would be desirable to identify the characteristics of auditory cues whose effectiveness is minimally degraded by noise and jamming.

3.0 FACILITY DESCRIPTION

3.1 Facility Overview

The purpose of the experimental facility is to serve as a test bed for evaluating the effectiveness of simulated auditory localization (SAL) as a method of providing directional information in head-coupled control/display systems. The basic task presented to the human test subject is to indicate the apparent location of a sound source based on SAL cues provided to the subject by way of headphones. The SAL cues change in real time as the subject's head moves, just as the proximal acoustic stimulus, at the entrance to the ear canal, changes in normal localization. The subject's head position is measured and used in real time to synthesize the SAL cues, which are fed back to the subject via headphones.

A general schematic of the apparatus for monitoring head position and for simulating localization cues is shown in Figure 19. Rotational and translational movement of the subject's head is measured at 60 Hz and used to switch an audio signal among 36 loudspeakers surrounding a model of the human head and ears in another room. Broad-band, high-fidelity microphones are placed at the entrances to the ear canals in the model head. The sound at the model head is amplified and fed back to the subject via headphones to produce SAL cues.

The room containing the model head and ears is covered with sound-absorbent material. The 36 loudspeakers are spaced at 10° intervals. The model head remains stationary, and the audio signal is transitioned between loudspeakers so as to simulate smooth, continuous movement of the signal relative to the head.

The following subsections provide more detailed descriptions of the SAL facility hardware, software, and experimental space.

3.2 Hardware Design

The functional diagram of the hardware for the SAL system is depicted in Figure 20. An IBM personal computer (PC) controls a number of peripheral devices attached through two parallel interfaces. The peripheral hardware required to implement the SAL system can be regarded as three major subsystems. The first of these is an audio production and control system,

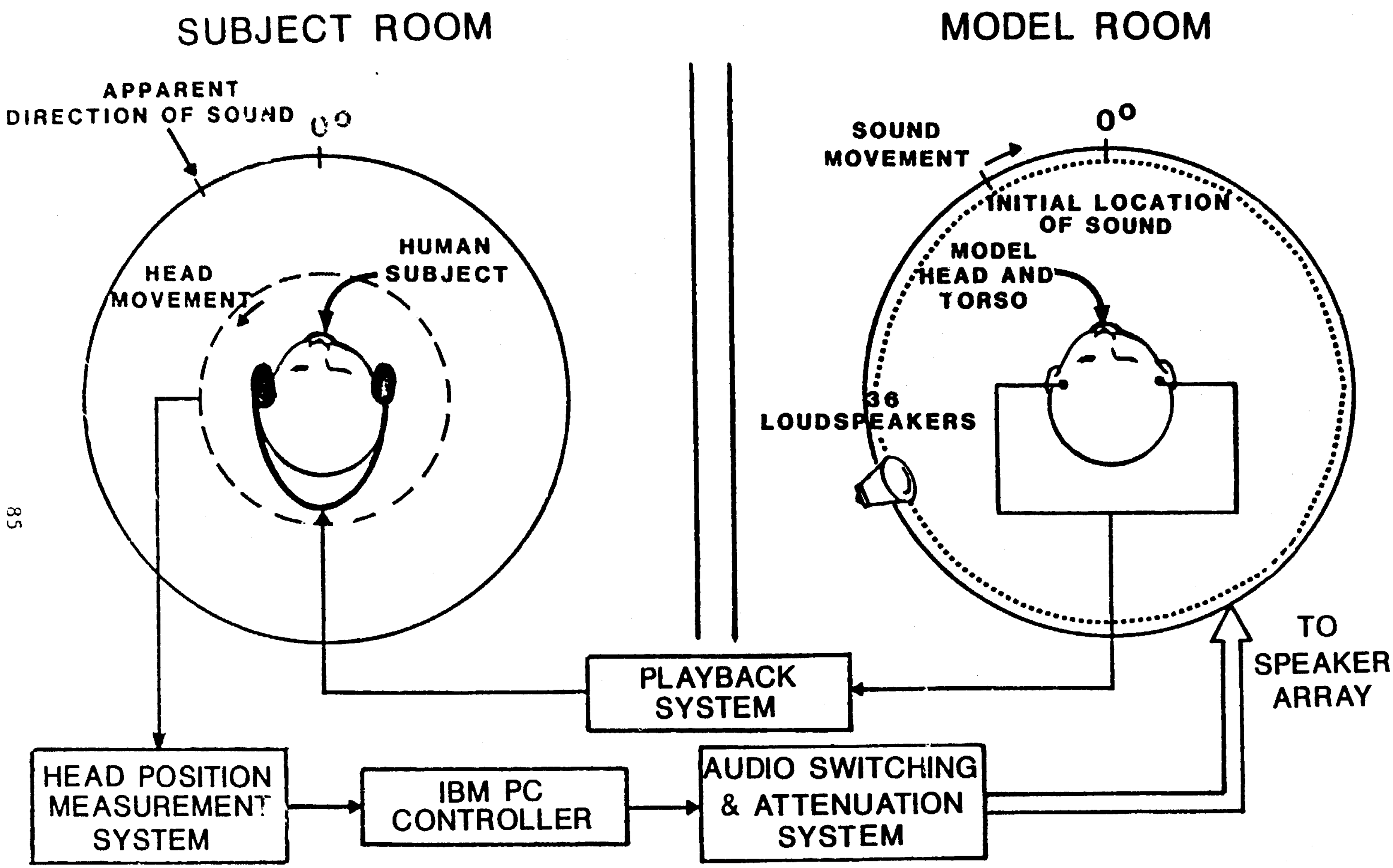


Figure 19. Schematic of SAL experimental facility.

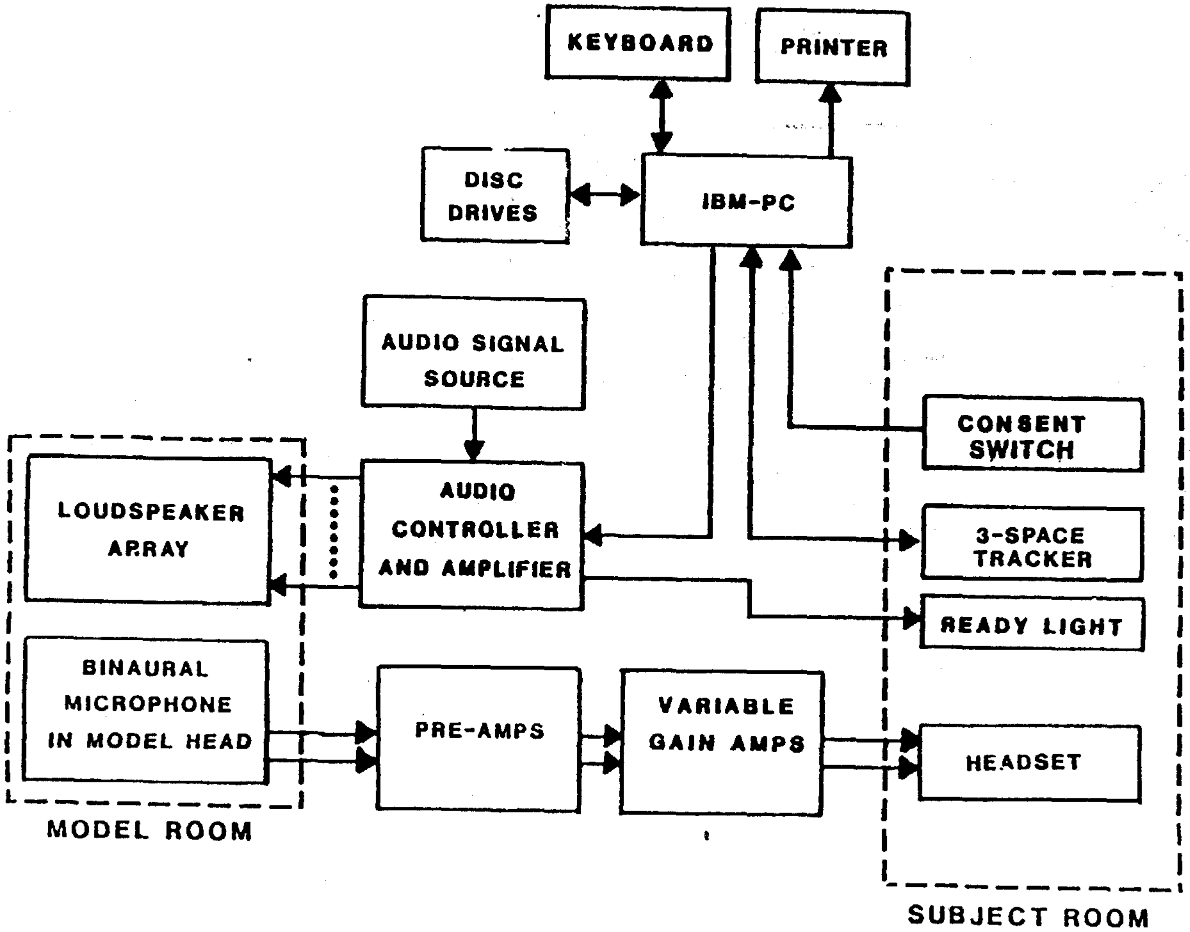


Figure 20. Functional diagram of the SAL system.

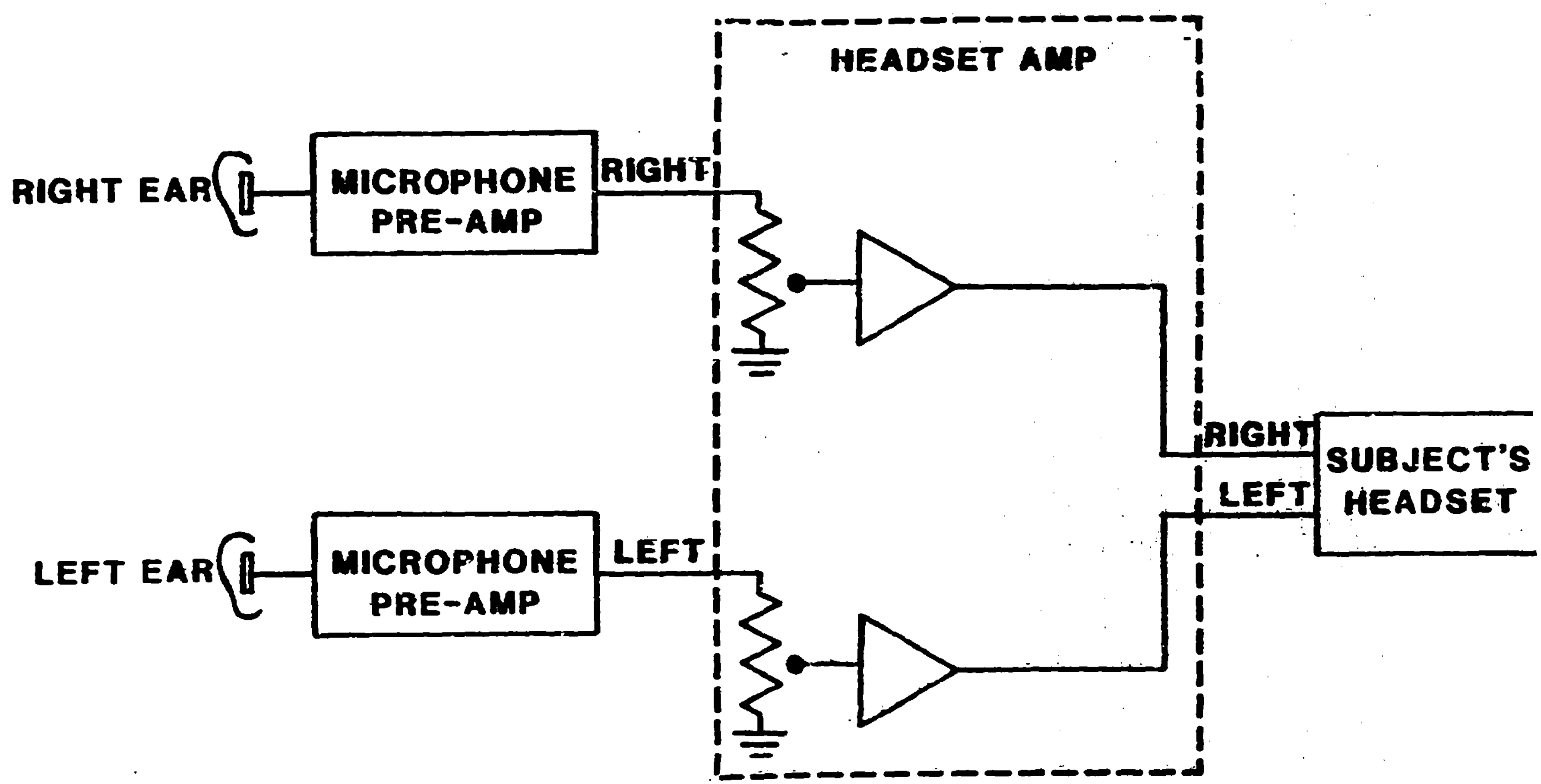
made up of an audio source, an audio controller with power amplifier, and a loudspeaker array. A second is the playback system, including the binaural microphones in the manikin's head, a two-stage amplification system, and a headset. The third subsystem is the 3-Space Tracker system, which provides head and hand position measurements.

The playback system, depicted in Figure 21, provides the audio feedback path from the speaker array to the subject. Two omnidirectional microphones are mounted in the ears of the manikin, located in the center of the speaker array. The audio signal is preamplified by two 30 dB amplifiers located at the base of the manikin. The signal then passes into the control room where the right and left audio signals are fed into two independent volume controls before final amplification. The outputs of the amplifiers are connected to the headset worn by the subject. The headphones are Sennheiser model HD230's, which are mounted in a Gentex HGO-26/P flight helmet with the visor removed.

The heart of the audio production and control system is the audio controller and power amplifier. The controller/amplifier, shown in Figure 22, controls the location of the sound source in the array of loudspeakers. Up to three audio signals can be mixed together to produce an audio output with a gain of 0 dB. The mixed output is then connected to three digitally-controlled step attenuators. After attenuation, the three audio signals are amplified and connected to the array of 36 loudspeakers by the relay control matrix.

The controller/amplifier is designed so that only three relays may be closed at any one time. Each of the three relays in turn is connected to a single amplifier. The level of attenuation applied to each amplifier-relay pair is controlled independently through the computer interface. An external subject ready light is also controlled through the computer interface.

The apparent position of the sound source relative to the manikin is controlled by activating combinations of loudspeakers at various attenuation values. Since the loudspeakers are located at 10° intervals around the manikin, sound sources at such positions (called "real" sources) are generated by simply activating the appropriate loudspeaker at minimum attenuation (maximum intensity) level. Sound sources located between 10° intervals (called "phantom" sources) are simulated by simultaneously activating a pair of adjacent loudspeakers. For example, a source at 5 degrees azimuth is


HEADSET AMP
RIGHT EAR
MICROPHONE PRE-AMP
RIGHT
LEFT EAR
MICROPHONE PRE-AMP
LEFT
RIGHT
LEFT
SUBJECT'S HEADSET

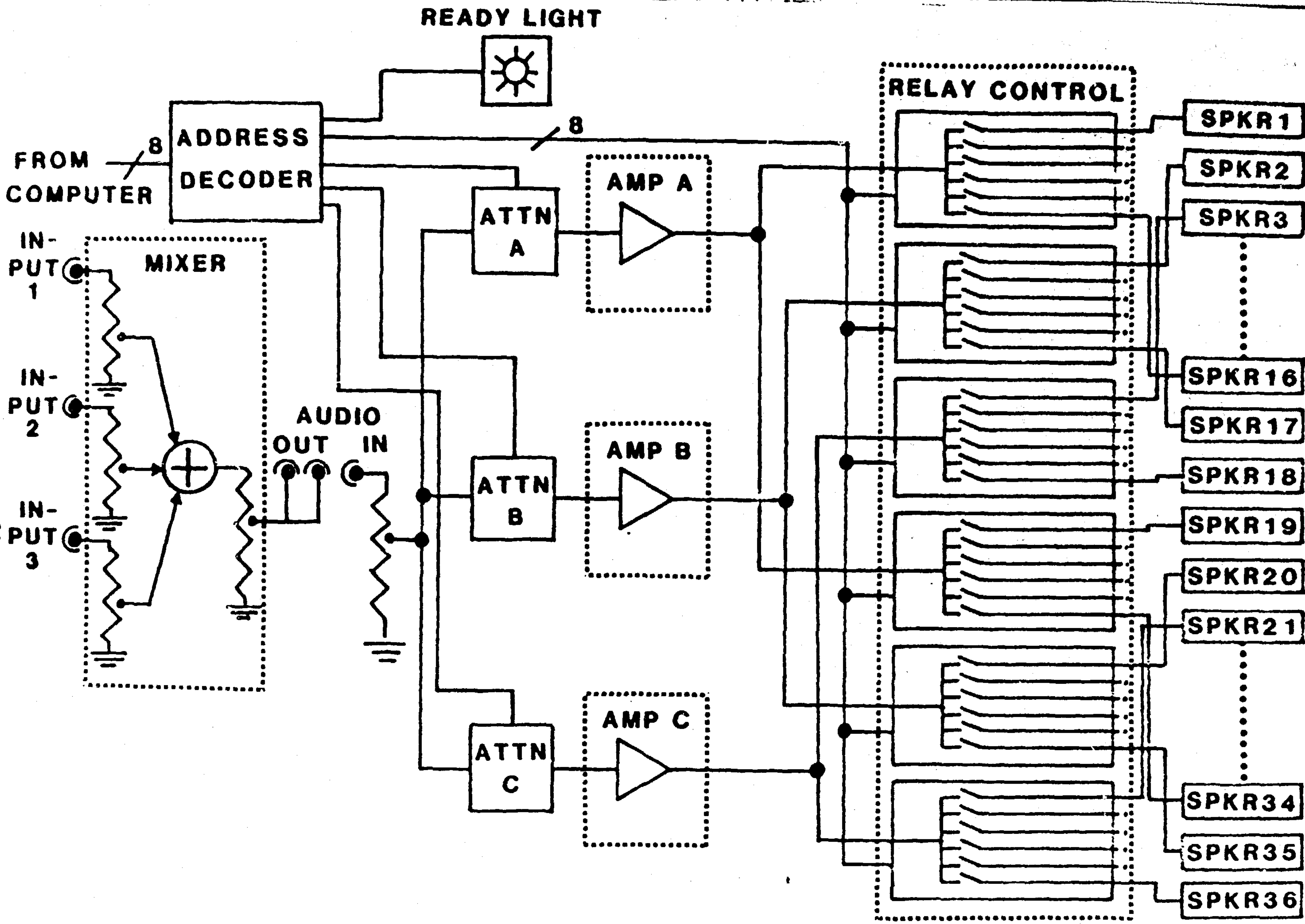


Figure 22. Audio controller and power amplifier.

simulated by activating the loudspeakers at both 0 degrees and 10 degrees azimuth at equal intensities (attenuation values). In order that phantom sources not sound louder than real sources, the two loudspeakers which create the perception of a phantom source between them are each played at reduced power. The relative power delivered to the adjacent loudspeakers determines the perceived position of the phantom sound source. The development of the algorithm for controlling apparent source position was undertaken as a part of the validation of the SAL facility, and is described in Section 4.4.

Head-position measurement is accomplished by using the 3-Space-Tracker system, model 3ST002, manufactured by Polhemus Navigation Sciences, Inc. The system utilizes a low frequency, magnetic field to determine the position and orientation of a sensor in relation to a source. The system has a maximum resolving power of 0.1 angular degree. However, its accuracy may be degraded when the source or sensor is in close proximity (6 feet) to large metallic objects. The system can use one or two sources and up to four sensors. Each source-sensor pair is uniquely identified in the system as a "station." The system can measure the relative position and orientation of up to four stations at the same time.

The 3-Space-Tracker system can be interfaced to the host system through a parallel or serial port; however, only the parallel interface insures the integrity of transmitted data through a hand-shaking protocol. A customized record of information representing position and orientation for all active stations can be sent to a host computer or other data collecton system continuously or at selected intervals.

In the present research, one source and two sensors are configured as two stations. The source is secured to a nonmetallic support attached to the ceiling in the center of the testing room. Metallic objects have been cleared from a 6 foot square area around source and sensor to reduce interference (see Section 3.4). One sensor is mounted on the helmet worn by the test subject. A second sensor is attached to a response indicator assembly held by the subject throughout experimental testing.

3.3 Software Architecture

3.3.1 Major Components

Figure 23 shows the digital hardware and software configurations for the SAL system in block-diagram form. The principal hardware components are the IBM PC with associated special purpose cards (monochrome adapter, memory expansion, expanded input/output capability), the audio controller/amplifier hardware, and the 3-Space-Tracker system. The principal software components depicted in Figure 23 are the executive control routine, the assembly language interface, and the speaker look-up table. The executive control routine, assembly language interface, and speaker look-up table are all resident in memory during program execution. The double and single headed arrows depict the information transfer between software and hardware components in the system and show the number of channels used in each function.

The executive control routine, written in Pascal, handles four major functions. The functions include interaction with the user, program initialization, experimental control and data collection, and calculation of the new sound position relative to the manikin (see Section 3.3.3).

The assembly language interface handles time-critical low-level communication and control operations, under the direction of the executive control routine. Handling these functions in assembly language helps to maximize the speed of these real-time control activities.

The speaker look-up table identifies the relays which must be closed and the values of each of the three attenuators in order to produce a sound source at a given azimuth relative to the manikin. The new sound azimuth is the address in the table at which the relay and attenuator device codes and values are stored. The relay and attenuator device codes and values are output to the audio controller/amplifier by the assembly language interface.

3.3.2 Overview of Software from User's Perspective

Control software for the SAL project was designed to be flexible enough to run all the anticipated experiments. The software handles the initialization of all devices and warns the experimenter if data storage space is low. The experimenter can set values for all experimental parameters and store them in the Master Test file. Table 3 shows the input parameters of the Master Test file.

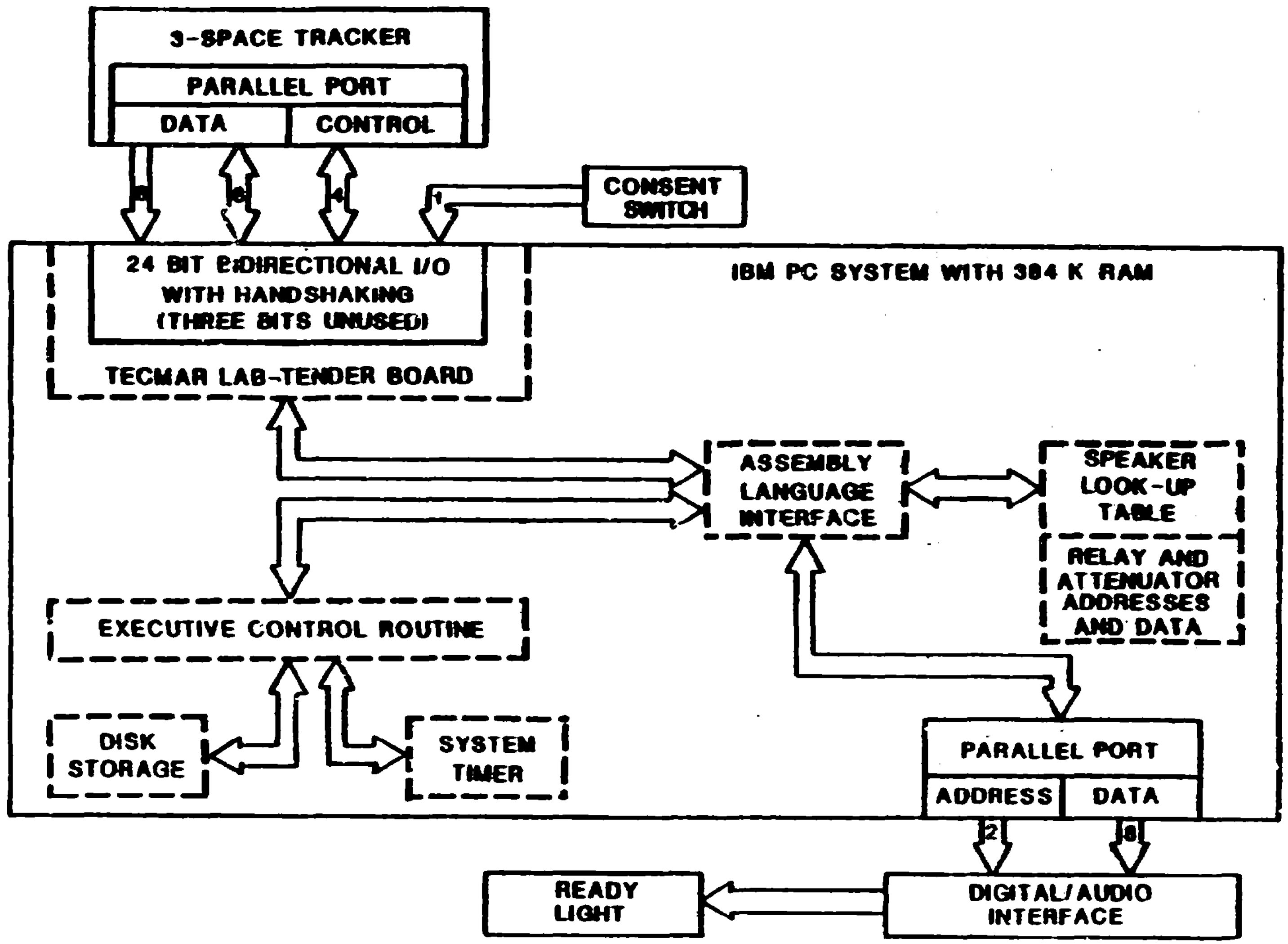


Figure 23. Block diagram of digital control logic and devices.

TABLE 3

Input Parameters in the Master Test File

Number of trials
Collect data (yes or no)
Auto or manual running mode
Nose-on time criterion (.1 sec)
Nose-on angle criterion (degrees)
Delay to start trial (sec)
Maximum sound duration (sec)
Intertrial interval (sec)
Monitor Finger positions,
Head position or Both
Initial Azimuth file name

The Initial Azimuth file contains the initial azimuth of the sound on each test trial, assuming that the subject is facing the 0 degree azimuth position. The initial azimuth is also the apparent azimuth of the sound throughout the test trial. The control software checks for the existence of this file and prompts the experimenter to create it if none is found by that name.

The starting block number and starting trial number can also be set at the beginning of a test session. This feature proves valuable when an interruption of testing requires restarting testing at some trial other than trial number one or some block other than block one. The experimenter can also choose to balance the left and right headphone channels at the beginning of the testing block, which is necessary when the audio stimulus source is changed.

At the beginning of each test block, the experimenter is prompted to enter the data filename, subject identification, experimental condition and time (date of testing). Data filenames are checked before the experiment starts to ensure no duplicate names are used. The control software warns the experimenter that the existing data file will be overwritten unless a new filename is specified. The Master Test file parameters collected on each block are written as the first line of each data file to serve as a permanent reference to identify the file later.

In addition to managing the presentation of azimuth locations for each trial, the control software also reports any errors detected in the data from the 3-Space tracking system and allows a trial to be aborted and restarted if difficulties occur during the course of a trial. At the end of a testing block all files are closed. After each test block, the experimenter is prompted to enter another Master Test file if an additional block of trials is to be run.

The events transpiring on each test trial from the experimenter's and subject's perspectives are as follows. First, the "ready" light at 0 degrees azimuth is illuminated, indicating that the SAL system is ready. The subject then presses the consent switch on the response indicator to indicate that he/she is ready. The response indicator is a hand-held, pistol-shaped object. A short delay then ensues (delay to start trial) before the system begins to monitor the subject's head position. The subject is required to be facing 0 degrees azimuth ("nose-on") for a minimum amount of time with a minimum accuracy before the trial continues. Once the nose-on criteria is met, the timer is started and the sound is turned on. The sound remains on until the consent switch is pressed a second time. At that point three events happen in close succession: (1) the elapsed time is recorded, (2) the position and orientation of the listener's head and/or the position of the hand-held response indicator is recorded, and (3) the sound is turned off. The data and identifying information for the trial are then written to disk. If there are further test trials to be run in the test block, the intertrial interval then ensues before the ready light is illuminated for the next trial. After the last trial in each block, the ready light flashes three times to signal the subject that the block has been completed.

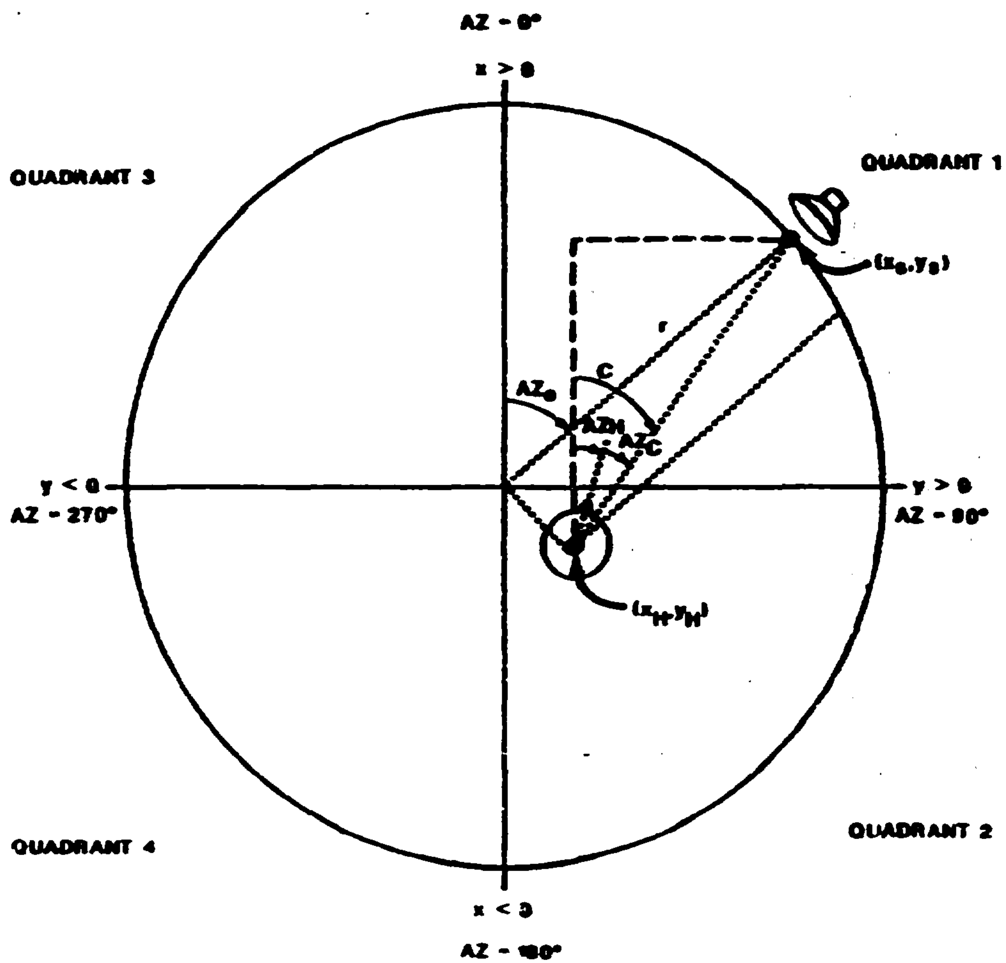
During testing, the system displays certain information on the PC display in order to allow the experimenter to monitor the subject's performance and detect any system problems. The data is displayed in near real time and includes the trial number, the apparent azimuth of the sound, the indicated azimuth of the sound, and the response time.

3.3.3 Real Time Calculation of Sound Position

The azimuth of the sound relative to the manikin was adjusted in real time to compensate for the listener's head movements. The objective of these adjustments was to make the sound appear to the listener to emanate from a stationary source at a specified azimuth for each test trial, called AZ_0. The listener's head movements were measured by the 3-Space Tracker system. As discussed earlier, the listener was seated in the "Subject" room, and localized sounds were presented via headphones. The manikin was located in the adjacent "Model" room, at the center of the circular array of 36 loudspeakers. The following calculations were made at a rate of 60 Hz, the sampling rate of the 3-Space Tracker system.

Three numbers regarding the listener's head orientation were input to the PC at 60 Hz:

> x_H, y_H = the coordinates of the center of the listener's head relative to the center of the loudspeaker array, measured as shown in Figure 24 (in inches).


AZ - 0°
x > 0
QUADRANT 3
QUADRANT 1
(x_s, y_s)
r
C
AZ_s
AZH
AZC
y < 0
AZ - 270°
y > 0
AZ - 90°
(x_H, y_H)
QUADRANT 4
QUADRANT 2
x < 0
AZ - 180°

AZ_H = the azimuth of the center of the listener's head relative to the center of the array, measured as shown in Figure 24 (in degrees).

On any given test trial, there were two constants which entered into the calculations:

r = the radius of the array of loudspeakers (114 inches).

AZ_O = the azimuth from which the sound should appear to come for the listener (i.e., the "initial azimuth" discussed previously).

If at the start of the test trial, the center of the listener's head was at the point ($x = 0$, $y = 0$) and the head was oriented in the horizontal plane such that $AZ_H = 0$, then the sound was presented to the manikin at azimuth AZ_O.

Suppose now that the listener's head is moved to a point (x_H, y_H) and rotated in azimuth an amount AZ_H. In order to make the sound appear to the listener to still come from the same point in space (x_s, y_s), the sound must be moved to azimuth AZ_C relative to the manikin, as shown in Figure 24.

It is clear from Figure 24 that:

$$AZ_C = C - AZ_H \text{ when } C \geq AZ_H$$

$$\text{and } AZ_C = C - AZ_H + 360° \text{ when } C < AZ_H.$$

The angle C is calculated as follows:

$$C = \text{Arctan } [(y_s - y_H)/(x_s - x_H)] + k$$

Where

$$x_s = r \cos (AZ_O)$$

and $$y_s = r \sin (AZ_O).$$

The constant k = 180° when the quantity $(x_s - x_H) < 0$; k = 360° when $(x_s - x_H) \geq 0$ and $(y_s - y_H) < 0$; and k = 0 otherwise.

3.3.4 Calculation of Perceived Sound Direction

As indicated in the Master Test file (Table 3), the SAL system can be configured to measure the subject's finger position only, head position only, or both. The so-called "finger" position is actually the position of the response indicator which the subject holds in the hand.

Depending on the instructions given the subject, either head position or finger position, or both can be used as indicators of the perceived (indicated) sound direction. When the "monitor finger position" (F) option is selected, the indicated sound direction written to the data file is based on the position of the response indicator, as discussed below. When the "monitor head" position (H) option is selected, the indicated sound direction is based on the head azimuth and position. When the "monitor both" (B) option is activated, two indicated sound directions are written to the data record, one based on head azimuth and position and the other based on the position of the response indicator. In all cases, the indicated sound direction is based on head/response indicator position just after the consent switch is closed.

3.3.4.1 Calculation of Perceived Sound Direction Based on Finger Position

The 3-Space Tracker provided measurements of the position of the response indicator relative to the 3-Space Tracker source, which was positioned directly above the listener's head at the start of each block of test trials. The listener was instructed to press the consent switch when the response indicator was pointed directly toward the sound. The perceived direction of the sound was taken as the azimuth of the response indicator relative to the 3-Space Tracker source.

The perceived direction of the sound, AZ_r, can be computed from the Cartesian coordinates (x_r, y_r) of the response indicator in the horizontal plane, relative to the 3-Space Tracker source, as follows:

$$AZ_r = \text{Arctan}\,[y_r/x_r] + k$$

Where $k = 180°$ when $x_r < 0$; $k = 360°$ when $x_r \geq 0$ and $y_r < 0$; and $k = 0$ otherwise.

3.3.4.2 Calculation of Perceived Sound Direction Based on Head Orientation and Position

As mentioned above, the software provides an option for recording final head position (just after the consent switch is pressed) as an indicator of the perceived direction of the sound source. Of course, using final head position as an indication of perceived sound direction is only appropriate when the subject was instructed to use his or her head orientation to indicate sound direction.

If the subject's head is rotated so that he or she is looking exactly in the direction of the sound (that is, head azimuth, AZ_H, equals the initial azimuth of the sound, AZ_o) then the direction of the sound relative to the manikin, AZ_c, will be 0 degrees azimuth. Any rotation or translation of the subjects head from that point, such that he or she is no longer directly facing the sound, produces a change in AZ_c, as described in Section 3.3.3.

Assuming that the subject is instructed to use head orientation as an indication of sound direction, the deviation represented by AZ_c can be regarded as the error in the subject's perception of the sound direction. The actual (correct) azimuth of the sound is, of course, AZ_o. Then the perceived sound location indicated by head position, AZ_p, can be calculated as follows:

$$AZ_p = AZ_o - AZ_c.$$

3.3.5 Response Time Measurement

The timing resolution of the IBM PC internal time-of-day clock, configured from the factory, is 100 msec. However, a computer clock is just a counter, driven at a certain frequency. Therefore, more precise timing is possible by altering the frequency of the timing counter to produce near-millisecond accuracy timing. In the SAL software, the counter frequency is set to 1000 Hz, which produces millisecond resolution. A restriction in the use of the timing routine is that disk drives cannot be accessed during timing periods because the system is being interrupted more frequently than with the 100 msec resolution clock and this interferes with disk access. The original counter frequency is restored when the timing period has ended, in order to allow disk access.

3.4 Experimental Rooms

Figure 25 depicts the SAL experimental room layout. The manikin, fitted with microphones, was located in the model room, surrounded by the circular array of 36 speakers. The manikin was the KEMAR model DB4004 manufactured by Industrial Research Products, Inc. The manikin was fitted with molded rubber pinnae, model DB065 (right ear) and model DB066 (left ear). The microphones were two model BT-1759's, provided by Knowles Electronics, Inc. The microphones were positioned at the manikin's ear canal entrances. The interior of the manikin head was stuffed with loose fiberglass insulation.

The loudspeakers were located 2.89 m from the center of the manikin head at 10 degree intervals. Certain tests with human subjects were conducted in the model room rather than the subject room; hence similar provisions were made in both rooms. Referring again to Figure 25, note that a portion of the ceiling has been raised in both the model and subject testing room to clear any electromagnetic influences from the area in which the 3-Space Tracker system was used to monitor head and/or hand position. This precaution insured that the highest possible accuracy was obtained from the 3-Space Tracker system. Figure 25 also shows placement of a uniform, opaque curtain drawn in a circle around the testing area. The curtain serves to eliminate any visual cues which might bias the subject's judgement of sound source location.

Acoustical treatment was applied to surfaces in the model room to reduce ambient noise levels and reverberation. The treatment included 24 sound-absorbent panels, each 4 ft. high by 4 ft. wide, placed just outside the periphery of the loudspeaker array. The panels were four-inch thick Sonex foam. In addition, the raised portion of the ceiling in the model room was covered with fiberglass insulation, loosely draped to form a scalloped surface. The acoustical performance of the room is discussed in Section 4.1.

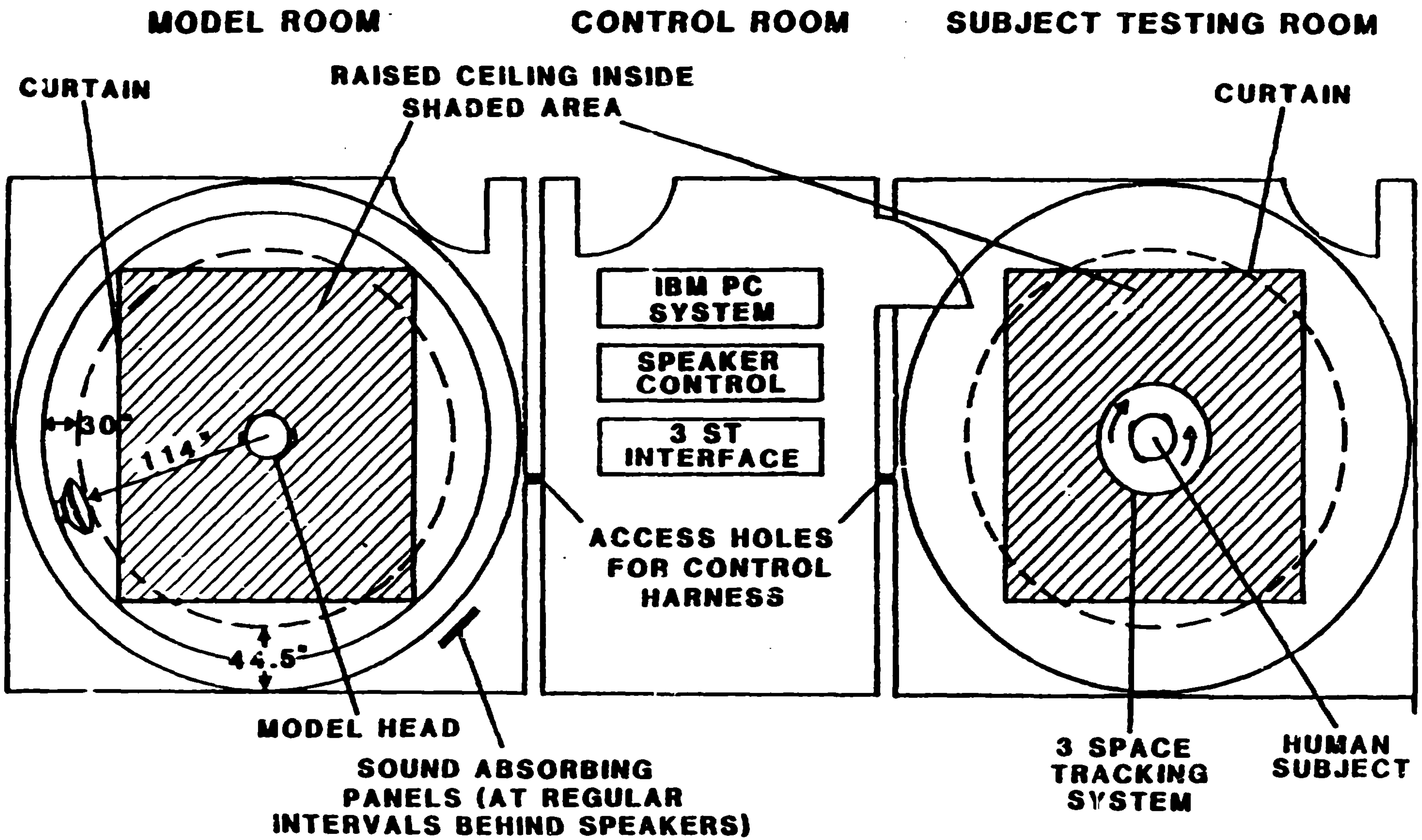


Figure 25. SAL experimental room layout.

4.0 VALIDATION OF THE EXPERIMENTAL FACILITY

In this section is described a series of tests which were conducted in order to evaluate the physical and psychological fidelity of the SAL facility. This section also covers refinements made to the SAL system, based on the results of these tests. The description of these procedures assumes a knowledge of the basic architecture of the SAL facility, which is presented in Section 3.0.

The tests and refinements dealt with five different topic areas which can affect the fidelity of the SAL system: (1) the acoustical character of the sound treated room in which SAL cues were generated, (2) the adequacy of the audio production system, including the power amplifiers and the loudspeaker array, (3) the frequency response of the audio playback system, from the manikin to the headphones worn by the listener, (4) the speed of the real-time hardware and software for controlling the direction of the sound relative to the manikin, and (5) formal experimental tests designed to evaluate the psychological fidelity of the SAL system.

4.1 Room Acoustics Tests

Brief acoustical tests were conducted to quantify the ambient noise level and reverberation time in the room where SAL cues were generated. This room contained the acoustic manikin and the array of 36 loudspeakers. The acoustical treatment of the room is described in Section 3.4.

Since it was anticipated that most of the experimental testing would take place on weekday afternoons, the test of ambient noise level was conducted on such an occasion. This provided some assurance that noise sources over which the experimenters had no control, such as traffic noise from nearby roads, would be at a level representative of that which would occur during experimental testing.

The ambient noise level was measured at the center of the room at the height of the manikin's ears (62 inches). The microphone was a Bruel and Kjaer type 4165 condenser microphone. The noise level was measured by a Bruel and Kjaer type 2203 precision sound pressure level meter.

The major noise sources within the building which could contribute to the noise level in the room included a central heating/cooling unit, a central blower, and a peripheral fan in the ventilation return duct immediately outside the room. The ambient noise test was repeated four times: once with all three sources on, with the fan off and the other sources on, with the fan and the blowers off and the heating/cooling unit on, and with all three sources off.

The measured ambient noise level with all three noise sources in the building on was 36 dBA. With only the peripheral fan off, the ambient noise was 33.5 dBA. In the remaining two conditions (central blower and peripheral fan off; and all three sources off) the measured noise level was 32 dBA.

The experimental testing was conducted with the peripheral fan off at all times. The heating/cooling unit and central blower operated intermittently during testing, under the control of an automatic thermostat. However, their operation made only a small difference in the ambient noise level (33.5 versus 32 dBA).

The reverberation decay time was measured for a wide-band (10 kHz) noise source. The source was a Bruel and Kjaer type 4205 sound power source. The noise source was placed midway between the center and one wall of the room. A microphone (same as that used in the ambient noise test) was again located in the center of the room at the height of the manikin's ears. The microphone signal was boosted by a Bruel and Kjaer type 2807 microphone preamplifier and fed to a Tectronix model 5223 digitizing oscilloscope. The noise source was operated at 92 dB re 1.0 pW for approximately 10 seconds in order to allow the noise to fill the room. The oscilloscope was triggered such that it displayed the steady-state noise signal just prior to its termination, and several seconds of the decaying noise after termination.

The display vertical dimension was in voltage units. The elapsed time for a 10 to 1 drop in voltage was recorded. Since voltage is proportional to sound pressure, p, and the Sound Pressure Level (SPL) is defined as 20 log (p/p_{ref}), a 10 to 1 drop in voltage is equivalent to a 20 dB drop in SPL. Thus 3 times the elapsed time required for a 10 to 1 drop in voltage is an estimate of the reverberation time, defined as the time required for a 60 dB drop in SPL. The reverberation time, so estimated, was approximately 350 msec for the broadband noise source.

4.2 Post-installation Tests of the Audio Production and Control System

The audio production and control system and its components underwent a series of four different tests and/or adjustments after the system was installed in the sound-treated room (see the preceding section of this report). The audio production and control system includes the following components: the audio source, the audio controller and amplification unit, and the 36 loudspeaker assemblies.

4.2.1 Loudspeaker Polarity Check

A simple polarity check was conducted to ensure that all the loudspeakers were radiating in phase. A 1.5 volt battery was connected to the loudspeaker terminals, with the positive terminal of the battery to the loudspeaker terminal marked positive. The direction of the movement of the woofer was noted. The loudspeaker terminals were reversed where necessary so that the woofer moved outward when the battery was applied as described.

4.2.2 Amplitude Response as a Function of Frequency

The purpose of this test was to determine the amplitude response over the audible frequency range of each of the 36 loudspeakers in the sound-treated room. This test served to identify any major differences between the loudspeaker units (including the crossover circuits) and any sources of unusual reverberation in the room.

A white noise source was played through a single loudspeaker and the amplitude response as a function of frequency from 0 Hz through 25 kHz was measured. The equipment included the microphone and preamplifier described in the last section, and a real-time digital spectrum analyzer (Hewlett-Packard model 3582A). The microphone was located at the center of the room. The test was repeated for each of the 36 loudspeakers.

A typical amplitude response is shown in Figure 26. The upper trace is the white noise input signal and the lower trace is the output of a loudspeaker installed in the sound-treated room. The vertical scale for the upper trace has been shifted downward so that the two traces can be displayed simultaneously. The overall difference in the amplitudes of the two traces is about 55 dBV. Notice that the loudspeaker response is virtually flat (within

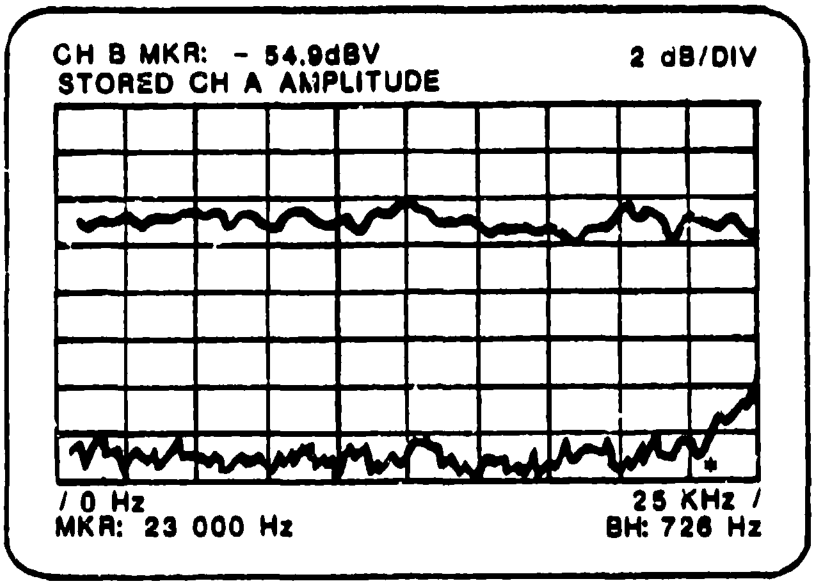


Figure 26. Input signal (upper trace) and amplitude response (lower trace) of the audio production and control system.

2 dBV) out to approximately 23 kHz. All 36 of the loudspeakers produced this same pattern, except that the locations of the small (less than 1 dBV) ripples varied from one loudspeaker to the next. The locations of the ripples in the white noise input signal also varied from one test to the next. This fact suggests that most of the variation in the loudspeaker responses is due to sampling and quantification error in the spectrum analyzer and/or variations in white noise source over time.

4.2.3 Equalization of the Loudspeakers

The purpose of this procedure was to remove any differences in the intensities of the loudspeakers for a constant signal applied. The SPL of each loudspeaker was measured with the audio control unit set at minimum attenuation value. The input signal was white noise at 100 dBA, measured from the mixer test point. The digital control unit and the mixer are described in Section 3.0.

Since the SPL of the loudspeakers differed slightly at minimum attenuation, the loudspeakers were equated to a constant SPL by increasing the attenuation values on the more intense units. As a consequence, a unique attenuator value was defined for each loudspeaker as its new minimum attenuation value. These values were used to adjust the speaker lookup table which controls the movement of the sound around the manikin, as described in Section 4.4.

4.2.4 Perceptual Equivalence of Phantom and Real Sources

The purpose of this test was to determine whether listeners could discriminate between sounds produced by a single loudspeaker (called a "real" source) and sounds produced by two loudspeakers activated simultaneously with the same total power (called a "phantom" source). In the SAL facility, loudspeakers are located at 10 degrees intervals in the horizontal plane surrounding the manikin or listener. Therefore, real sources can be played only at increments of 10 degrees. In order to produce a sound which appears to come from an intermediate azimuth, two loudspeakers must be activated simultaneously. The relative power emitted by the two loudspeakers determines the perceived location of the sound.

A white noise signal was played through "phantom" and real sources successively. An observer was seated at the center of the loudspeaker array with his/her ears at the same height as the center of the loudspeaker assemblies. The observer was asked to report whether sounds, played for 3 sec each, came from real or phantom sources. The room was darkened in order to eliminate visual cues. Two experienced observers were each run for about 30 trials each. Both observers performed at near chance level, i.e., they could not discriminate phantom from real sources under these conditions.

4.3 Tests of the Audio Playback System

Two types of tests were performed on the audio playback system. First, the SPL output of the left and right headphones was measured to determine whether the two units were of equal efficiency, and to determine whether their responses were proportional to the input intensity. A second test examined the amplitude response over the audible frequencies of the entire playback system, including the model room, the manikin, the microphones in the manikin's ears, the two-stage stereo amplification system, and the headphones. Since the manikin has pinnae, this chain of components should produce amplitude response functions over frequency like those produced by the human head and pinnae (i.e., a head-related transfer function or HRTF). Furthermore, the amplitude response functions should vary as a function of the direction of the source relative to the manikin in the same manner that human HRTFs vary as a function of source direction.

In the first test, the measurements were taken with the sound pressure level meter and microphone described in Section 4.1. The microphone was located in the center of, and approximately one inch from, the radiating surface of the headphone being tested. The headphones were removed from the helmet, and microphone and headphone were surrounded by fiberglass insulation material in order to reduce ambient noise inputs to the microphone as much as possible. The input signal was inserted after the second headphone amplifier, which is the same point from which the left and right headphone channels are equalized during experimental testing. The test was conducted with three different types of input signals and at three different signal intensities for each headphone.

The test conditions and results are shown in Table 4. It is clear from the results that the left and right channels are of equal efficiency, and that their responses are proportional to the input level. The response of the headphones is nearly identical for the white noise and the 2 kHz high-pass noise signals. As would be expected, the response is somewhat reduced for the 1 kHz low-pass noise signal.

In the second test, the microphone was positioned in front of the left or right headphone, and isolated from ambient noise as described for the first test. The microphone was boosted by a Bruel and Kjaer type 2807 preamplifier, and analyzed with a Hewlett-Packard model 3582A real-time digital spectrum analyzer. The manikin was positioned at the center of the model room, and one of three sources in the array of 36 loudspeakers was activated. The source activated was in the horizontal plane at 0, 90, or 270 degrees azimuth relative to the sagittal plane of the manikin. The input signal to the loudspeaker was white noise. The output level at the center of the room was 54 dBA. The signal levels of the left and right headphone channels, at the output of the second headphone amplifier, were set to 105 dBA. The amplitude response was measured in dBV relative to the input signal at intervals of 500 Hz, from 500 Hz to 20 kHz.

The results of the second test are shown in Figure 27. Three amplitude response functions are plotted for each channel of the playback system. The three functions are for a sound source at 0 degrees azimuth, for a source at 90 degrees azimuth on the same side as the active microphone (ipsilateral), or for a source at 90 degrees azimuth on the opposite side (contralateral). Up to 12 kHz, the functions are similar to the corresponding averaged HRTFs for human listeners shown in Figure 2. Since the present data were collected at intervals of 500 Hz, some of the fine structure typical of HRTFs does not appear. No comparisons can be made above 12 kHz since the averaged human HRTFs stop there, and little or no data on HRTFs have been collected above that frequency. The results suggest that the playback system accurately reproduces the transformation of the free-field sound by the head, torso, and pinnae.

TABLE 4

Sound Pressure Level Output (in dBA)
at the Left and Right Headphones
as a Function of Input Level and Type of Signal

Headphone channel	Input level (dBA)	Type of input signal: White noise	2 kHz high-pass noise	1 kHz low-pass noise
	90	54.5	55.0	54.0
left	105	68.5	68.5	65.0
	114	78.5	78.0	74.0
	90	54.5	55.0	54.0
right	105	68.5	68.5	65.0
	114	78.5	78.0	74.0

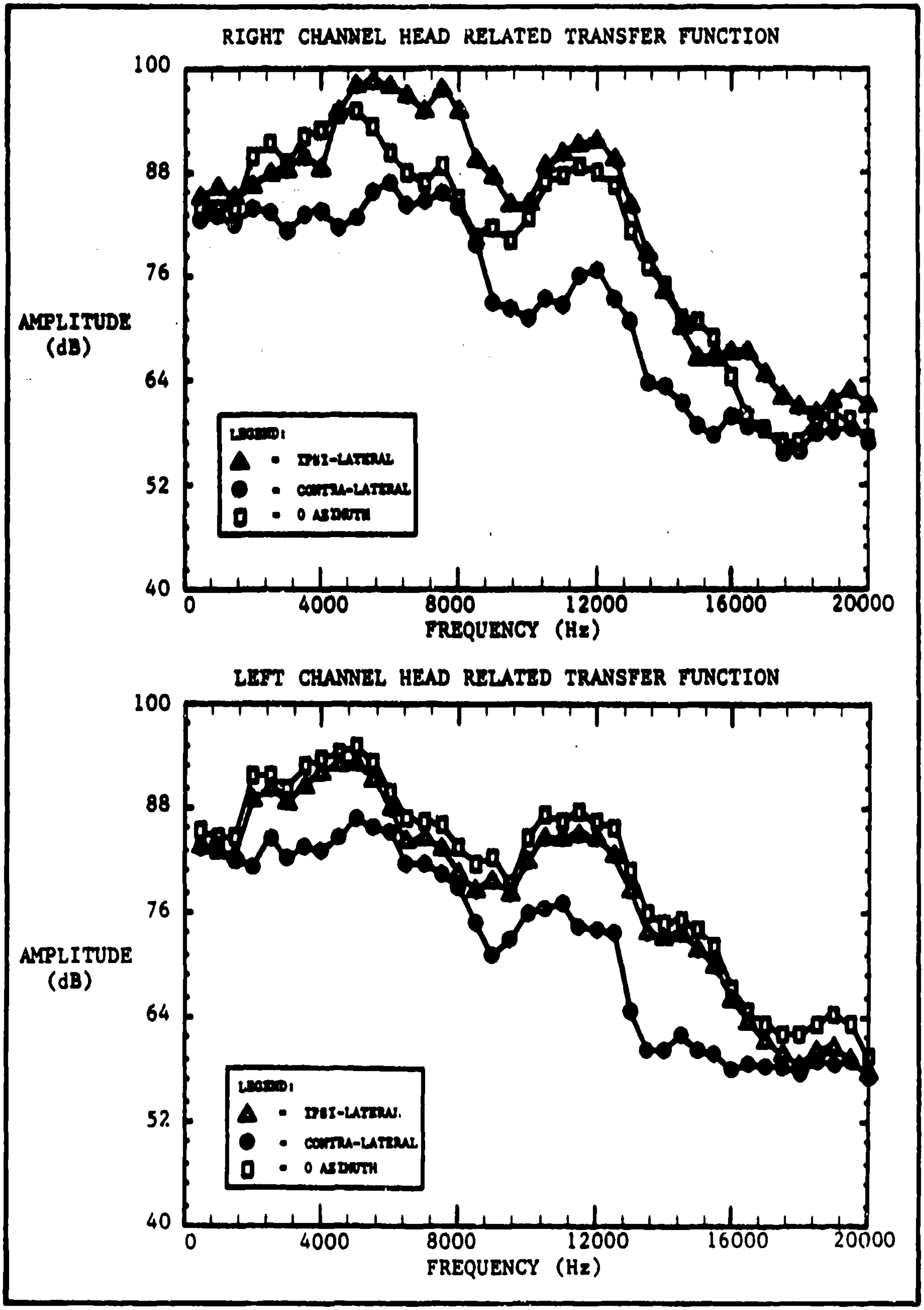


Figure 27. Amplitude response of the audio playback system.

4.4 Real-Time Control of Sound Direction

4.4.1 Interface with the 3-Space Tracker System

The sampling rate of the 3-Space tracking system is the rate limiting factor in correcting for the direction of sound when the subject moves his or her head. When only head position is measured (one station active) the sampling rate of the 3-Space tracking system is 60 Hz. Informal tests, where rapid head movements were made, demonstrated that the location of the sound remained stationary with a 60 Hz sampling rate. However, when finger and head position were measured at the same time (two stations active, where the sampling rate of the 3-Space tracking system was 30 Hz), correction of sound position lagged behind the rapid movement of the head. In effect, the sound did not remain stationary with both stations active. To eliminate this slow response in correcting sound location, a strategy of enabling only one station at a time was adopted.

At the beginning of each test trial the head station was enabled, so that the location of the sound could be corrected for head movements. The test subjects were instructed to identify the apparent direction of sound by pointing the response indicator in that direction and depressing the consent switch. When the switch closure was sensed by the PC, the finger station was enabled and the head station was disabled. After allowing settling time (100 msec) for the change in stations, the finger station was sampled twice and the second sample was kept. The finger station was then disabled and the head station was enabled to prepare for the next trial.

4.4.2 Determination of Relay and Attenuator Values

In order to keep the location of the sound stationary as the head moves, each report of head position by the 3-Space Tracker is followed by calculations in the PC executive routine to compensate for head translation, as described in Section 3.3.3. These calculations produce the new azimuthal location of the sound relative to the manikin, which is needed in order to make the sound appear stationary to the listener wearing headphones. The new azimuth value is used as an address to a table of 3600 sets of relay numbers and attenuator values, called the "speaker look-up table". Each entry in the table represents the relays and the attenuator values which produce a sound at a given one-tenth degree position in the loudspeaker array.

If the sampling rate of the 3-Space tracking system were fast enough to keep up with one-tenth degree changes in head position, then the corrected values obtained from the 3-Space Tracker could be used directly to move the sound around the manikin. Unfortunately, the maximum sampling rate of the 3-Space Tracker is too slow to keep up with one-tenth degree changes in head position during rapid head movements. During rapid head movements, the successive azimuths reported by the 3-Space Tracker can differ by several degrees. When the sound was moved around the manikin in increments of several degrees, the sound appeared to the listener to have an intermittent, "choppy" quality.

It was necessary, therefore, to develop a routine which plays the sound at intermediate azimuth values when two successive corrected values from the 3-Space Tracker differ by more than a threshold value. For example, if two successive values reported by the tracker differ by 2 degrees, the sound could be transitioned between the two positions in increments of, say, .5 degrees. Informal tests were conducted with experienced listeners to determine an increment size which produced the perception of a stationary, yet smooth, sound during the fastest head movements expected in the SAL system. It was found that moving the sound in .4 degree increments produced a perceptually smooth sound which appeared to remain stationary for reasonably rapid head movements.

The algorithm for transitioning the sound between adjacent loudspeakers as it moved around the manikin was also modified to further improve the smoothness of the sound during head movements. The original algorithm used a linear mixing of adjacent speaker voltage levels. Thus, to produce a source which appeared to come from midway between two loudspeakers, the adjacent loudspeakers were each set at 50 percent of their minimum attenuation values. This linear-voltage algorithm was based on the results of phasor analysis of stereophonic listening (cf., Bauer, 1961). However, the phasor analysis applies to only low frequency (<1500 Hz) sounds, since the auditory system uses phase information only in that region. For broadband sounds, it is possible that a more appropriate algorithm is a constant RMS voltage (i.e. constant total power) relationship. In this algorithm, the voltage applied to adjacent loudspeakers is proportional to the sine or cosine of the angle of the apparent source location relative to the real loudspeaker locations. When this algorithm was applied, the sound heard through the headphones appeared much smoother during head movement than it did with the linear-voltage rule. The sine-cosine algorithm was therefore incorporated into the speaker look-up table and governed the movement of sound around the manikin.

4.5 Psychological Fidelity Experiment

4.5.1 Introduction

An experiment was conducted in order to evaluate the psychological fidelity of the SAL apparatus, i.e., to quantify differences between localization in the SAL environment as opposed to normal free-field localization. Complete psychological fidelity would mean that the SAL apparatus is perceptually equivalent to normal localization, though not necessarily physically equivalent. The SAL apparatus simulates the physical conditions in normal localization by using a technique made up of three principal components.

First, sound sources at azimuths between whole-number multiples of 10 degrees are simulated by activating two adjacent loudspeakers simultaneously. Such sources are called "phantoms", as opposed to "real" sources, which represent a single loudspeaker and were located at whole-number multiples of 10 degrees. For any achievable physical set-up, there will always be some physical difference between the sound produced by two loudspeakers radiating simultaneously (i.e., a phantom) and that produced by a single loudspeaker. Fortunately, physical identity is not required; it is only necessary that phantom and real sources sound the same to the human listener.

Second, in the SAL facility the acoustic cues for localization are produced by radiating the KEMAR manikin, transducing the sounds at the entrance to the manikin's ear canals into electrical signals, and amplifying and reproducing the sounds through headphones. It is of interest to determine to what extent the manikin/playback system is perceptually equivalent to the listener's own ears during normal, free-field localization.

Third, in order to make a sound appear to be stationary for the listener wearing headphones in the SAL facility, the direction of the sound source relative to the manikin must be changed in real time to compensate for the listener's head movement. For example, as the listener rotates his or her head to the right, the sound impinging on the manikin must be moved an equal amount to the left. Of necessity, the manner in which the sound is moved around the manikin only approximates the manner in which a stationary sound moves relative to the head in normal free-field listening conditions. It is of interest to determine whether the method used to compensate for the listener's head movement in the SAL facility (see section 4.4.2) is the perceptual equivalent of a stationary source in free-field listening conditions.

In the psychological fidelity experiment, localization accuracy and response time were measured under conditions which used various combinations of the three components of the SAL simulation. In addition, a control condition was run which involved no simulation, i.e., this condition involved normal, free-field localization of a single (real) source. This allowed an evaluation of the overall fidelity of the SAL system and the identification of any of the three major components of the simulation responsible for departures from complete psychological fidelity.

4.5.2 Method

4.5.2.1 Experimental Design

Each of 4 listeners were tested in six conditions (3 source types by 2 listening conditions) represented in Table 5. A white noise stimulus was presented either to the unaided ears (normal, free-field conditions) or dichotically via headphones, with localization cues simulated by using the manikin. Three types of sources were used: real sources (at 10 degree intervals), phantom sources at 2 degrees from the nearest real source, and phantom sources at 5 degrees from the nearest real source.

For two listeners, the head remained stationary during localization. Another pair of listeners was encouraged to move their heads during localization and, in the simulated cues condition, the sound was moved about the manikin to compensate for the listener's head movements.

The dependent measures were localization accuracy (apparent source azimuth minus actual source azimuth) and response time. The response time was measured with millisecond accuracy, from the onset of the sound until the

TABLE 5

Mean Localization Error in the Psychological Fidelity Experiment as a Function of Type of Source and Head and Listening Conditions

Head condition	Listening condition	Subject number	Source Type: Two degree phantom	Source Type: Five degree phantom	Source Type: Real	Subject means	Group means
Fixed	Unaided	1	-1.585	0.187	-2.464	-1.287	-2.829
		2	-4.887	-3.334	-4.888	-4.370	
	Headphone	1	7.513	7.239	6.5[illegible]5	7.089	6.060
		2	4.964	5.127	5.000	5.030	
	Mean		1.501	2.305	1.041	1.616	
Moving	Unaided	3	3.201	3.717	2.045	2.988	3.993
		4	4.636	5.457	4.898	4.997	
	Headphone	3	-3.474	-3.725	-3.681	-3.627	-0.326
		4	2.862	3.058	3.004	2.975	
	Mean		1.806	2.127	1.56[illegible]	1.833	
Overall source means			1.654	2.216	1.304	1.725	

subject pressed the consent switch to indicate that he/she was pointing at the sound.

4.5.2.2 Procedure

In the fixed head condition, the listener was instructed to place his or her chin in a chin rest and to keep the head stationary. The chin rest was mounted on the platform which held the listener's chair. In both head conditions, the center of the subject's head was positioned directly under the 3-Space Tracker source. The chair was adjusted so that the listener assumed a comfortable posture.

In the unaided (free-field) listening conditions, the listener sat in the center of the loudspeaker array in place of the manikin. The listener was seated such that the entrance to his or her ear canals was at the same height as the center of the loudspeaker assemblies. The subject's chair was adjusted so that the center of the head was at the center of the loudspeaker array. In the simulated cues (headphone) conditions, the listener was seated in the ajoining "subject" room, and wore the helmet with the headphones. The helmet strap was fastened and adjusted so that the headphones were snug against the pinnae. The listener was positioned such that his or her head was in the same relative position occupied by the manikin's head in the model room.

In all testing, the listener's chair was surrounded by a circular, opaque curtain in order to block the view of the loudspeakers and/or corners of the room. The listener was seated at the center of the circle described by the curtain which had a radius of 78 inches. The room illumination was just bright enough to read comfortably, and the listener was not blindfolded.

In both experiments, the listener indicated the apparent direction of the sound by pointing the arm and hand, in which the response indicator was held. A magnetic sensor mounted on the indicator was used to measure its position in space as described in Section 3.3.4. The listener was instructed to point the indicator in the apparent direction of the sound as accurately as possible, and to press the consent switch, also mounted on the response indicator, while holding the response indicator steady. The subject was also told to use the right hand for sounds on the right side and the left hand for sounds on the left, and to hold the response indicator out at arm's length.

On each test trial, a sound was presented at one of 24 different azimuths, as shown in Table 6. Forty-eight test trials were run in each of 10 blocks, with brief rest periods between blocks. Each of the 24 azimuths was presented twice in each block in one of five pseudorandom orders. Five blocks with different pseudorandom orders were tested in the first test session for each listener. The listener returned for a second session, in which the same five blocks of azimuths were run in a different random order. Each listener was run in one practice/warm-up block of 10 test trials at the beginning of each session.

Since it would have been difficult to move the circular curtain from room to room between test sessions, listeners were tested in the model room in the first session and in the subject room in the second session. Thus all listeners ran in the unaided (free-field) condition in the first session and in the simulated cue (headphone) condition in the second session. Since the localization task is simple, there should be no additional improvement in performance after the warm-up trials, and therefore no differential practice effect for Session 1 versus Session 2. If this assumption is correct, performance in the unaided condition should not improve over test blocks.

All subjects were given a typed statement which described the purpose, nature, risks, and benefits of the experiment, and the experimenter also read the statement aloud to the subject. A standard audiometric test (ANSI 1969) was then administered. Subjects with greater than 30% hearing loss were not used in the experiments. The 30% criterion was computed by taking 30% of the difference between the threshold for young adults with normal hearing and the threshold for feeling for each frequency tested (cf., Sivian & White, 1933; Bekesy, 1960).

The headphone channels were balanced at 68 dBA at the beginning of each session which involved dichotic presentation. The signal level at the manikin's ears (or the listener's ears in the free-field listening conditions) was set to 54 dBA.

Each test trial began with the illumination of a light at the 0 degree azimuth position at the height of the subject's ears. The light indicated that the system was ready, and the subject was instructed to press the consent switch when he/she was ready for the trial to begin. The subject was also instructed to turn his/her head so that he/she was looking directly at the

TABLE 6

Azimuths at which Sounds were Presented in the Psychological Fidelity Experiment

Region (Degrees)	Side	Type of Source: Real	2 Degree Phantom	5 Degree Phantom
0	L	0	358	355
	R	0	2	5
45/315	L	320	318	315
	R	40	42	45
90/270	L	270	272	275
	R	90	88	85
135/225	L	230	228	225
	R	130	132	135

ready light. The system then sampled the subject's head azimuth until two successive azimuths within ± 5 degrees azimuth were obtained at least 0.5 sec apart in time. The light remained on until the subject met the head orientation criterion.

When the head orientation criterion had been met, the response timer was started and the sound was turned on. The sound remained on until the consent switch was pressed again. At that time, the PC automatically read the response timer, measured the position of the response indicator, and computed the apparent direction of the sound. The sound was then turned off, and the apparent sound direction and the response time were written to disk. Before the next trial began, an intertrial of 1.0 sec was imposed.

4.5.2.3 Subjects

The subjects were three male employees of the Georgia Tech Research Institute and one female who was not an employee. Subjects' ages ranged from 23 to 43 years. Subjects were paid for their participation.

4.5.3 Results

The data records of the psychological fidelity experiment included three numbers of interest for each test trial: (1) the actual azimuth at which the sound was presented, (2) the perceived (indicated) azimuth of the sound, based on the subject's pointing response, and (3) the response time. Localization accuracy was computed by subtracting the actual sound azimuth from the indicated sound azimuth. If the difference was greater than 180 degrees, then 360 degrees was subtracted from it to form the localization accuracy. If the difference was less than -180 degrees, 360 degrees was added to form the localization accuracy. These adjustments prevented misinterpretation of test trials on which the actual sound azimuth and the indicated azimuth were on opposite sides of the discontinuity at 360 degrees azimuth.

Examination of the raw data and debriefing of the subjects revealed that various conditions rendered the data erroneous on a small number of test trials. The first condition was a malfunction in the re-initialization of the 3-Space Tracker just before finger position was sampled. This caused exactly the same indicated azimuth to be recorded on two successive trials. The second of the two trials was therefore excluded from the data analysis in such case. A second condition involved the subject accidentally pressing the consent switch in the process of moving his or her arm to point toward the perceived direction of the sound. This condition produced large localization errors. In order to remove as many of such cases as possible from the data, test trials on which localization error exceeded 60 degrees were excluded from the data analysis.

On a few trials, the subject pressed the consent switch to extinguish the ready light, and the ready light did not go out immediately (because the subject had not met the "nose-on criterion"). The fact that the light did not go out caused the subject to believe that the switch had not been pressed firmly enough; hence the subject pressed the switch a second time. On a few trials, the nose-on criterion had been met and the sound had come on by the time the second switch closure was sensed. This produced very short response times (0.1 to 0.3 sec.). Examination of the raw data showed that valid response times were always greater than 0.7 sec. Therefore, test trials for which reported response time was less than 0.5 sec were excluded from the data.

Table 7 shows the number of test trials remaining after trials meeting any one or more of the above criteria were excluded. The number of excluded trials was 15 out of a total of 1920 observations, i.e., less than 1 percent of the trials. In this table, trials on which the sound was presented at azimuths of 355, 358, and 360 degrees have been pooled with trials for which the sound was presented at 0, 2, and 5 degrees.

Three types of summary statistics were computed for valid trials: (1) the root-mean square (RMS) localization error, (2) the arithmetic mean localization error, and (3) the mean response time. Mean localization error reflects the direction (left vs. right) of average perceived sound azimuth relative to the actual sound azimuth. Table 5 shows mean error as a function of type of source and listening conditions. Mean localization errors less than zero indicate that the average perceived sound azimuth was less (to the left of) the actual azimuth. In the head fixed condition, unaided localization appears to have been biased slightly to the left, while localization via headphones (SAL cues) was biased to the right. In the moving head condition, however, unaided localization was biased more to the right than was headphone localization. These effects probably reflect individual

TABLE 7

Number of Valid Trials in the Psychological Fidelity Experiment as a Function of Sound Azimuth and Head and Listening Conditions

Head condition	Listening condition	Subject number	Azimuth region (degrees) 225-230	270-275	315-320	355-005	040-045	085-090	130-135	Totals
Fixed	Unaided	1	29	30	30	60	30	30	30	239
		2	29	30	30	60	30	30	30	239
	Headphone	1	28	30	30	60	30	30	30	238
		2	29	30	30	60	30	30	30	239
Moving	Unaided	3	28	30	29	60	30	30	30	237
		4	28	30	30	60	30	30	30	238
	Headphone	3	29	30	30	60	30	30	30	239
		4	29	30	30	59	30	29	29	236
		Totals	229	240	239	479	240	239	239	1905

differences in perception and/or response strategy. In terms of absolute magnitude of bias, the simulated cues condition with head movement was comparable to the unaided condition with head movements. Note that mean localization error is virtually identical for all three types of sources (2 degree phantoms, 5 degree phantoms, and real sources).

The RMS localization error provides a measure of the variability or dispersion of perceived sound directions about the actual sound azimuths. The RMS error reflects two sources of variability: (1) the dispersion of perceived azimuths about the mean perceived azimuth, and (2) the deviation of the mean perceived azimuth from the actual sound azimuth.

Table 8 shows RMS localization error as a function of type of source and listening conditions. For three of the four subjects, localization error was slightly less in the unaided (normal) listening condition than in the headphones (simulated cues) condition. For the fourth subject this trend was reversed, i.e., he made greater localization errors in the unaided condition than with simulated cues. Excluding the fourth subject (for reasons explained below), overall localization accuracy was slightly better in the unaided condition (RMS error = 9.466 degrees) than with simulated cues (RMS error = 10.721 degrees).

Table 8 also shows a clear difference in RMS localization error for the fixed versus moving head conditions, but only small differences as a function of type of source. Thus 2 degree and 5 degree "phantom" sources were localized nearly as accurately as "real" sources.

The data of Table 8 are also suggestive of an interaction of small magnitude between head condition and listening condition. Again excluding subject 4, the difference in RMS error for unaided versus headphone listening when the head is allowed to move (2.285 degrees) is slightly greater than that in the fixed head condition (0.741 degrees).

Mean response time is shown as a function of type of source and listening condition in Table 9. As was the case with RMS error, the difference between mean response times for unaided versus headphone localization is relatively small for three of the four subjects. For subject number 4, response time was much faster in the unaided condition than with headphones (1.7 sec versus 5.1 sec). It is notable that this subject produced relatively high RMS

TABLE 8

RMS Localization Error in the Psychological Fidelity Experiment as a Function of Type of Source and Head and Listening Conditions

Head condition	Listening condition	Subject number	Source type: Two degree phantom	Five degree phantom	Real	Subject means	Group means
Fixed	Unaided	1	10.007	9.553	9.181	9.580	12.379
		2	15.714	14.293	15.526	15.178	
	Headphone	1	10.915	10.281	9.084	10.093	13.120
		2	16.910	16.429	15.098	16.146	
	Mean		13.387	12.639	12.222	12.749	
Moving	Unaided	3	3.684	4.215	3.017	3.639	7.574
		4	11.052	11.175	12.299	11.509	
	Headphone	3	5.708	6.104	5.961	5.924	6.218
		4	6.161	7.084	6.291	6.512	
	Mean		6.651	7.145	6.892	6.896	
Overall source type means			10.019	9.892	9.557	9.823	

TABLE 9

Mean Response Time in the Psychological Fidelity Experiment as a Function of Type of Source and Head and Listening Conditions

Head condition	Listening condition	Subject number	Source type: Two degree phantom	Five degree phantom	Real	Subject means	Group means
Fixed	Unaided	1	2.341	2.363	2.420	2.375	2.535
		2	2.774	2.694	2.613	2.694	
	Headphone	1	1.726	1.780	1.777	1.761	2.095
		2	2.444	2.512	2.328	2.428	
	Mean		2.321	2.337	2.285	2.315	
Moving	Unaided	3	4.515	5.118	4.736	4.790	3.231
		4	1.641	1.765	1.608	1.671	
	Headphone	3	6.198	6.033	6.198	6.143	5.602
		4	4.993	5.499	4.691	5.061	
	Mean		4.337	4.604	4.308	4.416	
Overall source means			3.329	3.471	3.297	3.366	

localization error in this condition, suggesting a speed-accuracy trade-off. Excluding this subject's data, the headphone condition produced slightly greater response times overall than did the unaided condition (the difference was 0.9 sec).

With the exception of subject number 4 in the unaided condition, there is a clear difference in response time for the head fixed versus head moving condition, with the latter being greater. As was the case with RMS error, there appears to be little or no difference in response time as a function of type of source (phantom versus real).

Again, the data suggests an interaction between head condition and listening condition. In the unaided condition (excluding subject 4), head movement lengthens response times by 2.3 sec relative to fixed head localization. In the headphone condition, however, head movement lengthens response time by 4.0 sec relative to the head fixed condition.

Tables 10 through 12 show the three measures localization accuracy and speed as a function of the actual azimuth of the sound source. Note in Table 10 for subjects 1, 2, and 4 that mean localization error tends to be negative for sound sources located on the subject's right side, and positive for sources on the subject's left. In other words, these subjects tended to exhibit a bias (systematic error) toward the front (0 azimuth) position.

Table 11 shows RMS localization error as a function of the actual azimuth of the sound. As expected, localization error increases with angular distance from the frontal sagittal plane in the fixed head condition. The increase in error with absolute azimuth appears to be comparable for the unaided and headphone conditions. For the head-fixed condition, the average difference in RMS error from the ± 135 degree region to the 0 degree region was 17.0 degrees for the unaided listening versus 16.0 degrees for headphone listening.

Head movement appears to eliminate the increase in RMS localization error as a function of absolute azimuth (see the lower half of Table 11). However, another effect now appears. RMS localization error is generally greater on the subject's left than on the right side. The effect is present in both the unaided and headphone listening conditions. A comparison of Tables 10 and 11 suggests that this effect is due to greater systematic biases in responding to sounds on the left. In other words, the larger RMS errors on the left side

TABLE 10

Mean Localization Error in the Psychological Fidelity Experiment as a Function of Actual Source Azimuth and Head and Listening Conditions

Head condition	Listening condition	Subject number	Azimuth region (degrees) 225–230	270–275	315–320	355–005	040–045	085–090	130–135	Means
Fixed	Unaided	1	8.400	4.925	8.251	-0.328	-11.983	-2.522	-16.727	-1.289
		2	14.492	1.082	5.609	0.876	-13.014	-11.101	-33.681	-4.358
	Headphone	1	21.004	3.132	10.216	4.720	0.311	8.132	5.544	7.222
		2	29.778	13.909	7.970	6.293	4.473	0.017	-27.482	5.156
	Mean		18.419	5.762	8.012	2.890	-5.053	-1.369	-18.087	1.683
Moving	Unaided	3	2.005	3.450	3.214	3.844	3.778	2.666	0.867	2.959
		4	14.794	19.418	12.840	3.696	-4.776	-4.115	-5.906	4.956
	Headphone	3	-7.509	-8.638	-8.071	-3.039	-0.037	1.142	-0.172	-3.670
		4	8.930	7.298	7.906	2.911	-5.593	-1.357	0.126	2.892
	Mean		4.555	5.382	3.972	1.853	-1.655	-0.416	-1.271	1.784

TABLE 11

RMS Localization Error in the Psychological Fidelity Experiment as a Function of Sound Azimuth and Head and Listening Conditions

Head condition	Listening condition	Subject number	Azimuth region (degrees) 225-230	270-275	315-320	355-005	040-045	085-090	130-135
Fixed	Unaided	1	9.423	5.500	8.442	1.371	12.046	8.364	16.829
		2	15.009	5.320	6.116	2.098	13.218	12.500	33.768
	Headphone	1	21.391	5.445	10.517	4.816	2.037	8.779	7.458
		2	30.049	14.116	8.821	6.476	5.203	2.798	27.756
	Mean		18.968	7.595	8.474	3.691	8.126	8.110	21.453
Moving	Unaided	3	2.313	3.572	3.482	3.984	4.072	3.116	1.377
		4	14.957	19.733	12.973	4.015	4.855	5.245	6.099
	Headphone	3	7.841	8.821	8.320	3.456	2.847	2.855	2.327
		4	9.502	7.481	8.167	3.507	5.926	2.231	2.914
	Mean		8.653	9.902	8.236	3.741	4.425	3.362	3.179

TABLE 12

Mean Response Time in the Psychological Fidelity Experiment as a Function of Sound Azimuth and Head and Listening Conditions

Head condition	Listening condition	Subject number	Azimuth region (degrees) 225–230	270–275	315–320	355–005	040–045	085–090	130–135	Means
Fixed	Unaided	1	2.896	2.405	2.131	2.205	1.723	2.563	2.764	2.362
		2	3.178	3.012	2.508	2.293	2.253	2.730	3.218	2.686
	Headphone	1	2.070	1.935	1.679	1.548	1.446	1.703	2.149	1.760
		2	2.710	2.480	2.065	2.242	2.820	2.288	2.581	2.429
	Mean		2.714	2.458	2.096	2.072	2.061	2.321	2.678	2.309
Moving	Unaided	3	4.824	4.964	4.831	4.514	4.639	4.666	5.430	4.798
		4	2.258	1.761	1.664	1.186	1.590	1.796	2.020	1.693
	Headphone	3	7.004	6.848	6.361	5.509	5.734	6.304	5.863	6.142
		4	4.878	5.641	5.796	3.059	6.524	5.953	5.886	5.100
	Mean		4.741	4.804	4.663	3.567	4.622	4.680	4.800	4.431

are not due to greater variability of perceived sound azimuths on the left, but to a systematic bias to perceive sources on the left as displaced from their actual azimuth.

Mean response time is shown as a function of azimuth in Table 12. The increase in response time from the 0 degree region to the ±135 degree region was greater for the head moving condition (1.2 sec) than for the fixed head condition (0.6 sec). There also appears to be a small interaction of listening condition by azimuth. The increase from center to extreme azimuths was about 0.6 sec for headphone listening, versus 0.1 sec for unaided listening (excluding subject 4).

Tables 13 and 14 show RMS error and response time for the 10 blocks test trials distributed over two days. The data have been categorized by actual sound azimuth; those in front (±45 degrees) being shown separately from lateral azimuths (35 to 135 and 225 to 275 degrees). Since sources at the side and back are more difficult to localize, at least with the head fixed, it was anticipated that they would exhibit a greater practice effect. However, this was not borne out. From Table 13, it can be seen that there is little or no evidence of an improvement in localization accuracy over blocks for either the frontal or the lateral azimuths in the unaided listening condition. However, in the headphone listening condition there is a tendency for RMS localization error to decrease over blocks. Excluding subject 4, the overall mean RMS error on the last test block of headphone listening (7.6 degrees) is virtually identical to that last block of unaided listening (7.5 degrees). This is notable because the headphone condition produced slightly greater RMS error than unaided listening when the results were averaged over test blocks (see earlier discussion).

It is also notable that RMS error on the first block of headphone (simulated cue) localization (block 6) is roughly equal to the RMS error throughout all five blocks of the unaided condition (blocks 1 to 5) excluding subject 4. These data suggest that there was neither a positive nor a negative transfer effect in going from the normal localization cues in the unaided condition to the simulated cues in the headphone condition.

Table 14 shows mean response time for the 10 test blocks over two days of practice. There is a tendency for response time to decrease over test blocks in both the unaided and headphone conditions. The decrease in response time

TABLE 13

RMS Localization Error in the Psychological Fidelity Experiment as a Function of Listening Condition, Blocks of Practice, Sound Azimuth, and Head Condition

Head condition	Subject number	Sound position	Unaided Blocks					Headphone Blocks				
			1	2	3	4	5	6	7	8	9	10
Fixed	1	± 45°	8.214	7.082	6.978	7.743	7.363	6.755	6.322	6.934	6.716	5.849
		Others	10.765	11.045	10.629	12.279	11.698	14.556	13.068	12.372	13.139	10.418
	2	± 45°	8.366	7.496	8.294	6.745	7.469	7.650	7.001	7.188	6.941	6.998
		Others	18.464	21.389	20.123	21.060	19.093	19.961	20.753	23.295	21.410	22.939
	Mean	± 45°	8.290	7.284	7.636	7.244	7.416	7.203	6.662	7.061	6.829	6.424
		Others	14.615	16.217	15.376	16.670	15.396	17.259	16.911	17.834	17.275	16.679
Moving	3	± 45°	4.320	4.974	3.456	3.814	4.046	6.131	6.617	6.061	4.615	2.618
		Others	3.537	3.256	2.426	3.162	3.136	7.992	5.921	6.688	6.269	4.690
	4	± 45°	6.549	6.590	7.868	8.717	8.524	6.370	5.358	6.287	5.243	6.550
		Others	7.690	10.325	15.995	17.642	17.419	7.152	7.387	6.039	4.828	8.981
	Mean	± 45°	5.435	5.782	5.662	6.266	6.285	6.251	5.988	6.174	4.929	4.584
		Others	5.614	6.791	9.211	10.402	10.278	7.572	6.654	6.364	5.549	6.836
Grand mean		± 45°	6.862	6.533	6.649	6.755	6.851	6.727	6.325	6.618	5.879	5.504
		Others	10.114	11.504	12.293	13.536	12.837	12.415	11.782	12.099	11.412	11.757

TABLE 14

Mean Response Time in the Psychological Fidelity Experiment as a Function of Listening Condition, Blocks of Practice, Sound Azimuth, and Head Condition

Head condition	Subject number	Sound position	Unaided Blocks 1	2	3	4	5	Headphone Blocks 6	7	8	9	10
Fixed	1	± 45°	2.023	2.120	2.377	2.138	1.743	1.513	1.361	1.578	1.765	1.567
		Others	2.291	2.672	3.193	2.547	2.642	1.892	1.971	2.022	1.948	1.997
	2	± 45°	3.048	2.689	2.394	1.922	1.662	2.647	3.158	1.950	1.958	2.002
		Others	4.328	2.962	2.902	2.659	2.771	2.885	2.739	2.457	2.209	2.265
	Mean	± 45°	2.536	2.405	2.386	2.030	1.703	2.080	2.260	1.764	1.862	1.785
		Others	3.160	2.817	2.998	2.603	2.707	2.394	2.355	2.240	2.079	2.131
Moving	3	± 45°	4.321	4.624	5.196	4.975	3.959	5.754	5.655	6.574	5.619	5.291
		Others	5.065	4.801	5.067	5.782	4.221	6.948	6.799	6.691	6.834	5.364
	4	± 45°	1.665	1.426	1.309	1.373	1.232	4.836	5.660	4.894	3.978	3.261
		Others	2.107	1.946	1.975	1.579	1.709	5.462	6.346	5.547	5.417	5.248
	Mean	± 45°	2.993	3.025	3.253	3.174	2.546	5.295	5.658	5.734	4.799	4.276
		Others	3.586	3.374	3.521	3.881	2.965	6.155	6.573	6.119	6.126	5.306
Grand mean		± 45°	2.764	2.715	2.819	2.602	2.124	3.688	3.959	3.749	3.330	3.030
		Others	3.373	3.095	3.259	3.242	2.836	4.274	4.464	4.179	4.102	3.719

over test blocks is about the same for both listening conditions and for both head conditions. On the last test block of each day (blocks 5 and 10) mean response time was roughly 1.0 sec faster in the unaided condition as compared to the headphone condition.

4.5.4 Discussion and Conclusions

The results make possible an evaluation of the three principal components of the SAL facility, as discussed earlier (see Section 4.5.1). The major results are summarized in Table 15.

The algorithm for simulating sound sources at azimuths which fall between the locations of the loudspeakers was successful. The difference in RMS localization error for "phantom" as opposed to real (actual) sources was less than 0.5 degree. In addition, RMS localization accuracy was comparable for phantoms at 2 degrees and at 5 degrees from the nearest real source.

The SAL equipment for synthesizing the acoustic cues necessary for localization was also successful. When the listener's head remained stationary, localization accuracy with simulated cues was nearly as good as in normal, unaided listening conditions. With the listener's head fixed, the RMS localization erors for simulated and normal cues were 13.1 and 12.4 degrees, respectively. The differences in mean reponse time for these two conditions were less than 0.5 sec.

The third major component of the SAL system was the method for moving the sound around the manikin in real time in order to compensate for the listener's head movements. The results for subject 3 in Table 13 illustrate this effect. In the first test block with simulated cues, RMS error was 7.1 degrees, versus 3.9 degrees for the first block of unaided localization. By the fifth test block, localization accuracy was virtually identical for simulated cues (RMS = 3.7 degrees) and normal cues (RMS = 3.6 degrees).

In summary, the SAL facility produces a high fidelity simulation of normal, free-field localization. The initial difference in localization accuracy for simulated cues delivered via headphones versus unaided localization was about 3 degrees RMS. This difference decreased to virtually zero after one session (5 blocks of 48 test trials). The overall difference in mean response time with simulated as opposed to normal cues was 1.0 sec, and this difference was still present after one session of practice.

TABLE 15

Summary of Major Results of the Psychological Fidelity Experiment

- PHANTOM vs. REAL SOURCES
  - DIFFERENCE IN RMS LOCALIZATION ERROR < 0.5°
  - NO DIFFERENCES FOR 2° vs. 5° PHANTOMS
  - MEAN RT's VIRTUALLY IDENTICAL
- UNAIDED vs. HEADPHONE LISTENING (WITH HEAD FIXED):
  - DIFFERENCE IN RMS LOCALIZATION ERROR $\doteq$ 0.7°
  - DIFFERENCE IN MEAN RT < 0.5 SEC
- UNAIDED vs. HEADPHONE LISTENING (HEAD MOVING)
  - RMS LOCALIZATION ERROR
    - TEST BLOCK 1: 3.9° vs. 7.1°
    - TEST BLOCK 5: 3.6° vs. 3.7°
  - MEAN RT
    - TEST BLOCK 1: 4.7 SEC vs. 6.3 SEC
    - TEST BLOCK 5: 4.0 SEC vs. 5.3 SEC

It is likely that response times for the simulated cues condition would have been shorter if localization accuracy had been stressed less heavily in the instructions. That is, listeners can probably trade accuracy for speed.

Further research is needed in at least three areas related to SAL. First, research should be undertaken to optimize the fidelity of SAL cues. The results of the present research suggest that the real-time alteration of the sound with head movement is the area where there is greatest potential for improvement. Second, further study is needed to evaluate the effects of practice on localization performance with simulated cues, and the form of the speed-accuracy tradeoff in localization. For example, what is the impact on localization accuracy of introducing incentives for localizing more quickly? In a task in which SAL cues are used to direct visual attention, what is the impact on visual search time of speed incentives?

Third, a coordinated program of psychoacoustic research and engineering development is needed to develop an electronic synthesizer of directional audio signals which is suitable for airborne applications.

5.0 EXPERIMENTAL RESEARCH

5.1 Introduction

This research was directed toward identifying how the characteristics of auditory signals affect their usefulness as directional cues in the cockpit. Two experiments were conducted in which listeners localized sounds on the basis of simulated cues delivered via headphones. The character of the stimulus in the temporal and frequency domains was manipulated. The dependent measures were localization accuracy and response time. In the first experiment, the frequency composition of the stimulus was varied by using high-pass and low-pass noise stimuli with various frequency cutoffs. In Experiment II, intermittent noise stimuli were used, and burst duration, rise time, and duty cycle were varied.

Both experiments involved repeated testing of experienced subjects, i.e., a repeated measures design. The stimuli were presented dichotically, via headphones with the localization cues generated by the manikin, as described in Section 3.0. Subjects were encouraged to move their heads while localizing the sounds, and the direction of the sound relative to the manikin was changed in real time and coordinated with head movement in order to make the sound appear to be stationary to the listener. Subjects indicated the apparent azimuth of the sound by pointing their arm and hand and pressing the consent switch on the response indicator, as described in Section 4.5.2 for the psychological fidelity experiment. As in the previous experiment, the subjects were screened for participation by using the ANSI 1969 audiometric test. Subjects with greater than 30% hearing loss or abnormally shaped pinnae were not used.

The experimental procedure was the same as that part of the psychological fidelity experiment in which the subject received simulated cues and was free to move the head.

5.2 Experiment I

5.2.1 Introduction and Method

The purpose of the first experiment was to determine the importance of various parts of the audible spectrum for localization performance with simulated cues. The stimuli included white noise and high- and low- pass filtered noise signals with steep skirts, as shown in Table 16. The cutoff frequencies of the stimuli were designed to pass portions of the spectrum associated with some localization cues while rejecting portions of the spectrum associated with others. Table 17 shows the stimuli and the associated localization cues, based on previous research (see Section 2.0 for a review of the research).

Each of four subjects ran in two sessions, in each of which they experienced nine blocks of test trials. Each block consisted of 40 test trials. On each test trial, the stimulus was presented at one of 8 azimuths in the horizontal plane: 20, 60, 100, 140, 220, 260, 300, or 340 degrees. Each of the 8 azimuths was presented 5 times in each block, in one of three different pseudorandom orders. Each of the 9 types of noise stimuli was presented on one block in each session, with the order of presentation of the noise stimuli randomized over sessions and subjects.

Each session lasted approximately two hours. Each test subject received an initial block of 10 practice/warm-up trials at the beginning of each session. A white noise stimulus was presented on practice/warm-up trials. Subjects were given a 2 to 3 minute rest period between blocks, during which time they could remove the headphones and move about the laboratory.

Thresholds were measured for each of the nine stimuli for each subject prior to the experiment by the method of adjustment. During threshold measurment, the stimuli were presented dichotically with the sound at 0 degrees azimuth relative to the manikin, and the sound direction was not corrected for head movement. The experiment was run with the intensity of each stimulus adjusted to 33 dBA sensation level for each subject.

The subjects were one female and three male employees of the Georgia Tech Research Institute who were paid for their participation. They ranged in age from 23 to 29 years.

5.2.2 Results

The raw data for Experiment I were first processed as discussed in Section 4.3.3 to correctly interpret cases in which the actual sound azimuth and the azimuth indicated by the subject were on opposite sides of the 360,0

TABLE 16
Filtered Noise Stimuli
Used in Experiment I

Type of filtering	-3dBV Cut-off frequency (kHz)	Slope of skirt (dBV/Octave)
High-pass	2	78.3
	4	69.6
Low-pass	1	-79.0
	2	-72.8
	4	-66.1
	8	-59.4
	12	-26.8
	16	-29.3

TABLE 17

Stimuli for Experiment I
and Localization Cues Available

Type of Noise	Cut-Off Frequency (kHz)	Cues Available			
		ITD	IAD	Torso	Pinnae
Low-Pass	1	✓			
	2	✓	✓		
	4	✓	✓	✓	
	8	✓	✓	✓	✓
	12	✓	✓	✓	✓
	16	✓	✓	✓	✓
White		✓	✓	✓	✓
High-Pass	2		✓	✓	✓
	4		✓		✓

degree discontinuity. The raw data were then further filtered to remove test trials on which any one or more of three conditions was met. The three conditions were response time less than 0.5 sec, localization error greater than 60 degrees, or exactly the same indicated azimuth reported as on the immediately preceding trial. The rationale for these exclusions is discussed in Section 4.3.3. Forty-five of 2,880 test trials, or 1.5 percent, were excluded from the data.

As in the psychological fidelity experiment, three summary measures were computed: mean localization error, root mean square (RMS) localization error about the actual sound azimuth, and mean response time. Each measure was computed for each of 576 cells representing all possible combinations of 9 stimulus types, 8 azimuths, 2 days of practice, and 4 subjects. The mean across subjects for each measure was then computed for each of 144 conditions shown in Table 18. Multiple, planned comparisons analysis of variance (ANOVA) was used to test hypotheses concerning the RMS error and response time data.

The planned comparisons were designed to contrast localization performance for conditions in which different localization cues were available (see Table 17). Additional comparisons were also made to examine the effect of stimulus bandwidth, practice (Day 1 versus Day 2), type of noise stimulus, and source azimuth. The complete set of comparisons is shown in Table 19.

The RMS localization error results are shown in Tables 18, 20, and 21. The overall mean RMS error is remarkably similar for the nine types of noise stimuli. The only statistically reliable contrast ($F(1,3) = 10.64$, $p < .05$) was that between white noise and 2 kHz high-pass noise (the "ITD" contrast). This result suggests that the presence of low frequencies (<2 kHz), which provide interaural time difference (ITD) cues, facilitates localization with simulated cues.

Given that the ITD contrast was significant, it is surprising that the "ITD + torso" contrast was not significant. In fact, the 4 kHz high-pass noise produced a smaller RMS error (i.e., more accurate localization) than did white noise. The inconsistency of these two contrasts and the variability in RMS error across noise types suggests that the significant ITD contrast may be spurious (i.e., a type I error).

TABLE 18

RMS Localization Error as a Function of
Stimulus Noise Type, Source Azimuth,
and Day in Experiment I

DAY	Noise type	Source Azimuth (degrees) 20	60	100	140	220	260	300	340
	1LP	8.873	7.205	5.722	5.044	6.898	7.141	8.474	8.705
	2LP	11.154	7.579	8.539	12.467	8.279	5.786	8.053	10.020
	4LP	9.149	4.948	3.384	6.273	9.199	8.777	10.368	11.100
	8LP	9.392	4.865	5.623	8.837	9.842	7.361	11.005	11.427
1	12LP	11.597	7.597	6.606	7.933	6.490	10.977	12.995	10.348
	16LP	12.797	5.984	5.918	6.621	10.342	9.274	14.011	11.623
	WN	7.852	3.983	6.394	8.730	8.091	9.819	7.903	11.569
	2HP	10.047	6.479	5.721	6.742	11.026	10.985	14.748	12.212
	4HP	10.081	5.353	4.312	5.690	8.126	8.059	11.524	10.963
	1LP	7.008	4.695	5.338	9.883	7.387	6.104	7.283	9.113
	2LP	3.861	4.335	4.972	7.740	5.863	5.771	5.398	7.014
	4LP	5.666	4.252	5.181	6.599	5.002	4.287	5.481	8.822
	8LP	4.989	3.975	7.075	7.726	5.173	5.643	5.707	9.512
2	12LP	6.412	4.073	6.264	5.953	5.139	5.771	7.384	8.849
	16LP	7.260	3.840	7.202	8.622	6.293	7.100	9.945	8.793
	WN	7.632	3.023	6.843	8.051	5.778	5.631	7.961	10.192
	2HP	7.027	4.908	5.773	11.077	6.727	8.913	10.236	11.113
	4HP	6.062	4.905	4.367	7.662	7.203	4.588	6.373	6.601

TABLE 19

Coefficients for
Planned-Comparison ANOVA's
in Experiment I

	Variable: Noise type								
	Low-pass cutoff (kHz)						White noise	High-pass cut-off (kHz)	
Contrast	1	2	4	8	12	16		2	4
IAD	+1	-1	0	0	0	0	0	0	0
IAD + torso	+1	0	-1	0	0	0	0	0	0
IAD, torso and pinnae	+1	0	0	-1	0	0	0	0	0
Bandwidth	0	0	0	+3	+1	-1	-3	0	0
ITD	0	0	0	0	0	0	+1	-1	0
ITD + torso	0	0	0	0	0	0	+1	0	-1
Torso w/o ITD	0	0	0	0	0	0	0	+1	-1
Pinnae	0	0	+1	-1	0	0	0	0	0
Torso w/ ITD	0	+1	-1	0	0	0	0	0	0

Variable: Source azimuth (degrees)

Contrast	20	60	100	140	220	260	300	340
Side	+1	+1	+1	+1	-1	-1	-1	-1
Linear	+3	+1	-1	-3	-3	-1	+1	+3

Additional contrasts included the simple effects of source azimuth (side) and source-azimuth (linear) for each noise type and for each day of practice.

TABLE 20

RMS Localization Error as a Function of Stimulus Noise Type and Day in Experiment I

Noise type	Day 1	Day 2	Means
1LP	7.258	7.101	7.179
2LP	8.985	5.619	7.302
4LP	7.900	5.661	6.781
8LP	8.544	6.225	7.384
12LP	9.318	6.231	7.774
16LP	9.571	7.382	8.476
WN	8.042	6.889	7.466
2HP	9.745	8.222	8.983
4HP	8.013	5.970	6.992
Means	8.597	6.589	7.593

TABLE 21

RMS Localization Error as a Function of
Source Azimuth and Day in
Experiment I

	Source azimuth (degrees)							
Day	20	60	100	140	220	260	300	340
1	10.105	5.999	5.802	7.593	8.699	8.686	11.009	10.885
2	6.213	4.223	5.890	8.146	6.063	5.979	7.308	8.890
Means	8.159	5.111	5.846	7.869	7.381	7.333	9.158	9.887

Referring again to Table 20, RMS error on Day 2 is consistently smaller than on Day 1. The Day effect approached significance ($F(1,3) = 5.13$, $p = .1084$). This result is consistent with the practice effect seen in the headphone condition of the psychological fidelity experiment. Recall that RMS localization error decreased over practice blocks for the headphone (simulated cues) condition. The fact that there was no decrease in RMS error in the unaided listening condition suggests that the improvement in localization accuracy is due to adaptation to the simulated cues rather than improvement of the pointing response.

Table 21 shows RMS error as a function of Day and Azimuth. There was a significant effect of the Azimuth-side contrast ($F(1,3) = 1.409$, $p<.05$), reflecting greater error in localizing sounds on the listener's left. Among the simple effect contrasts, there was a significant Azimuth-side effect on Day 1 ($F(1,3) = 23.49$, $p< .05$), but not on Day 2.

Table 22 shows mean localization error (systematic bias) as a function of Day and Azimuth. These data suggest that the Side effect noted above on Day 1 is due to systematic localization error. For sounds at azimuths of 220 to 340 degrees on Day 1, the listeners pointed an average of 7 to 10 degrees to the left of the actual sound position. This systematic error was greatly reduced on the second day. These results suggest that subjects may have been less accurate in pointing with their left hand than with their right. Examination of individual subject's data showed that the effect was present for all 4 subjects, one of whom was left-handed. Recall that subjects were instructed to point with the left hand when the sound occurred on the left side, and with the right hand for sounds on the right side.

The only other significant comparisons among the simple effects were the Azimuth-side effects for the following stimuli: 4 kHz, low-pass noise ($F(1,3) = 12.67$, $p < .05$); 16 kHz low-pass noise ($F(1,3) = 39.08$, $p < .05$); white noise ($F(1,3) = 18,52$, $p < .05$); and 2 kHz high-pass noise ($F = 31,45$, $p< .05$).

Tables 23 and 24 show mean localization response time as a function of Noise type, Azimuth, and Day. The only significant contrast was the effect of Day ($F(1,3) = 18.27$, $p < .05$). Mean response time averaged 0.8 sec greater on Day 1 than on Day 2.

TABLE 22

Mean Localization Error as a
Function of Source Azimuth and Day
in Experiment I

	Source azimuth (degrees)							
Day	20	60	100	140	220	260	300	340
1	-1.709	0.719	-1.157	-4.781	-7.326	-7.880	-10.080	-10.051
2	-0.848	1.867	-2.487	-3.480	-2.679	-4.513	-5.885	-6.691
Means	-1.278	1.293	-1.822	-4.131	-5.002	-6.196	-7.983	-8.371

TABLE 23

Mean Response Time as a Function of Stimulus Noise Type and Day for Experiment I

Noise type	Day 1	Day 2	Means
1LP	3.130	2.449	2.790
2LP	3.564	2.790	3.177
4LP	3.370	2.629	3.000
8LP	3.023	2.631	2.827
12LP	3.281	2.864	3.073
16LP	3.920	2.712	3.316
WN	3.602	2.515	3.059
2HP	3.496	2.493	2.994
4HP	3.370	2.671	3.021
Means	3.417	2.639	3.028

TABLE 24

Mean Response Time as a
Function of Source Azimuth and Day
for Experiment I

	Source azimuth (degrees)							
Day	20	60	100	140	220	260	300	340
1	3.893	2.996	3.111	3.339	3.892	3.251	3.135	3.722
2	2.707	2.515	2.485	2.775	2.977	2.649	2.475	2.532
Means	3.300	2.756	.798	3.057	3.435	2.950	2.905	3.127

5.3 Experiment II

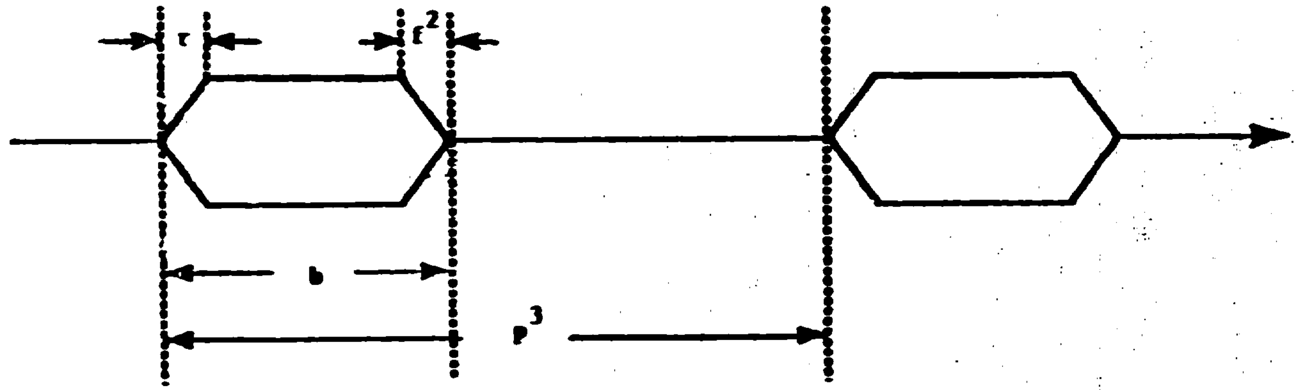
5.3.1 Introduction

The purpose of this experiment was to examine the effects on localization performance of rise time, repetition rate (duty cycle), and burst duration for trains of high- and low-frequency noise bursts. From studies of lateralization, it appears that the importance of rise time as a cue depends on signal duration. The longer a signal lasts, the more ongoing interaural time and amplitude differences appear to contribute to lateralization. As the signal duration approaches 300 msec, rise time no longer has an effect on the lateralization just-noticeable difference (jnd) (Mills, 1972; Durlach & Colburn, 1978).

As discussed in Section 2.0, the cues for horizontal plane localization differ at high and low frequency. At low frequency, ongoing ITD is the dominant cue. At high frequency several cues are available, including IAD, transient and envelope time differences, and monaural cues. Hence, it was of interest to study the trade-off between rise time and burst duration separately at high and low frequency.

The repetition rate variable is of interest primarily for high frequency stimuli. As noted in Section 2.0, the auditory system is sensitive to interaural time differences associated with the envelope of complex high-frequency sounds. The present experiment examined the effect of envelope frequency (i.e., repetition rate) on localization performance at both high and low frequencies.

A four-factor, completely crossed design was used. The four factors included: (1) two types of noise bursts (2 kHz high-pass and 1 kHz low-pass), (2) two rise times, (3) two burst durations, and (4) two duty cycles. The values of rise time, burst duration, and duty cycle used are shown in Figure 28. The combinations of duty cycle and burst duration employed resulted in three values of repetition rate, as shown in the right-most column of Figure 28. Notice that repetition rate is the interaction of duty cycle and burst duration. The design of this experiment allows an independent examination of the effects of duty cycle, burst duration, and repetition rate.



DUTY CYCLE(d)[1]	RISE TIME (r)	BURST DURATION (b)	REPETITION RATE
10%	1 msec	50 msec	2.0 Hz
10%	20 msec	50 msec	2.0 Hz
10%	1 msec	250 msec	0.4 Hz
10%	20 msec	250 msec	0.4 Hz
50%	1 msec	50 msec	10.0 Hz
50%	20 msec	50 msec	10.0 Hz
50%	1 msec	250 msec	2.0 Hz
50%	20 msec	250 msec	2.0 Hz

LEGEND:

[1] The duty cycle is defined as $[b/p] \times 100$

[2] The fall time was constant at 20 msec in all conditions

[3] The repetition rate, $f = [d/b] \times 10$

Figure 28. Parameters of the intermittent noise stimuli used in Experiment II.

As in Experiment I, there may be a main effect of noise type, reflecting the efficiency of the localization cues available with each stimulus (see Table 17). There should also be an effect of the laterality (i.e., absolute azimuth) of the sound source.

5.3.2 Method

Each of four subjects ran in two sessions, in each of which they experienced 16 blocks of test trials. Each block consisted of 30 test trials. On each test trial, the stimulus was presented at one of 6 apparent azimuths: 20, 60, 100, 260, 300 or 340 degrees. Each of the 6 azimuths was presented 5 times in each block, in one of 8 different pseudorandom orders. Each of the 16 stimuli, defined by combinations of Noise type, Rise time, Duty cycle, and Burst duration, was presented during one block in each test session. The order of presentation of the stimuli was randomized over subjects and sessions.

Each session lasted approximately two and one-half hours. Each subject received an initial block of 10 practice/warm-up trials at the beginning of each session. The stimulus for the practice block was the same as that for the first block of the session. Subjects were given a 2 to 3 minute rest period between blocks, during which time they could remove the headphones and move about the room.

Thresholds were determined by the method of adjustment for each of the 16 stimuli for two of the four subjects. For these two subjects, the differences in thresholds as a function of Rise time, Burst duration, and Duty cycle were quite consistent, as shown in Table 25. Therefore, for the remaining two subjects, the thresholds were measured for only four of the 16 stimuli, and thresholds for the remaining 12 stimuli were estimated. The thresholds for these 12 conditions were estimated by adding the "threshold difference" values shown in Table 25, which are based on the measured thresholds for the first two subjects.

Threshold measurements were taken prior to the first experimental session. During threshold measurement, the stimuli were presented dichotically with the sound at 0 degrees azimuth relative to the manikin, and the sound direction was not corrected for head movement. The experiment was run with the intensity of each stimulus adjusted to 22 dBA sensation level for each subject.

TABLE 25

Thresholds for Experiment II Stimuli

Noise type	Stimulus Characteristics: Rise time (msec)	Duty cycle (percent)	Burst duration (msec)	Thresholds (dBA): Subject No. 1	Subject No. 2	Mean threshold (dBA)	Threshold difference values used to estimate thresholds for S's 3 and 4[1]
1 kHz Low-pass	20	50	250	71.0	70.5	70.6	---
			50	70.5	70.5		
		10	250	70.5	70.5	70.3	0
			50	70.0	70.0		
	1	50	250	73.0	73.0	73.1	2.5
			50	73.0	73.5		
		10	250	78.0	78.0	77.8	7.0
			50	77.5	78.0		
2 kHz High-pass	20	50	250	65.0	66.0	64.1	---
			50	62.5	63.0		
		10	250	66.0	66.5	67.0	3.0
			50	67.5	68.0		
	1	50	250	70.5	72.5	70.9	7.0
			50	70.0	70.5		
		10	250	77.5	78.0	77.4	13.0
			50	76.0	78.0		

[1] See text for explanation

The subjects were three male and one female employees of the Georgia Tech Research Institute, who were paid for their participation. The subjects ranged in age from 19 to 43 years.

5.3.3 Results

As in the previous experiments, the raw data were processed and some trials were excluded to remove spurious data and account for the discontinuity at 360 degrees. Twenty four trials of the total 3,840, or 0.6 percent, were excluded from the analysis reported herein. Mean localization error, RMS localization error, and mean response time were computed for each subject in each of 96 cells representing combinations of 16 stimulus types by 6 source azimuths.

The RMS error and mean response time data were submitted to a repeated-measures ANOVA. The ANOVA model included the main effects of all four stimulus variables (Noise type, Rise time, Burst duration, and Duty cycle), the linear component of the azimuth variable, and all two and three-way interactions of the foregoing variables. The coefficients for the Azimuth-linear contrast were +1, 0, -1, -1, 0, and +1 for source azimuth of 20, 60, 100, 260, 300, and 340 degrees, respectively.

Table 26 shows mean localization error as a function of the stimulus variables and azimuth. Most of the means are quite small in absolute value, indicating that there was relatively little systematic error in localization. This result also means that the RMS error measure reflects primarily the dispersion of indicated azimuths about the cell means rather than the deviation of the cell means from the actual source azimuths.

The ANOVA for RMS error produced only three significant effects. The linear trend with absolute source azimuth was statistically reliable ($F(1,3) = 13.43$, $p < .05$). The last row in Table 27 shows that RMS localization error was smallest at ± 20 degrees and increased monotonically to ± 100 degrees azimuth. Even though the azimuth effect is significant, the effect is quite small. The RMS error increases an average of only about 1.0 degree from ± 20 to ± 100 degrees azimuth.

The three-way interaction of Burst duration by Duty cycle by Azimuth-linear approached significance ($F(1,3) = 6.28$, $p < .10$). Table 28 shows the

TABLE 26

Mean Localization Error as a Function of Stimulus Variables and Source Azimuth

Noise type	Rise time (msec)	Burst duration (msec)	Duty cycle (percent)	Source azimuth (degrees)					
				20	60	100	260	300	340
1 kHz low-pass	1	50	10	1.550	4.917	3.330	−0.453	2.064	0.208
			50	1.311	4.386	2.698	1.963	1.991	0.549
		250	10	1.835	6.173	5.501	−0.807	0.618	0.800
			50	1.503	3.803	2.888	−1.090	−0.805	0.014
	20	50	10	2.223	5.599	2.403	1.698	3.973	2.310
			50	1.482	3.595	3.655	1.581	2.872	−0.586
		250	10	1.816	4.387	3.192	0.502	1.857	0.706
			50	0.621	2.230	2.409	−0.788	−0.885	−2.171
2 kHz high-pass	1	50	10	1.500	4.454	4.987	−1.357	−0.418	−0.375
			50	2.040	2.980	4.605	−0.203	0.305	−0.142
		250	10	−0.473	1.919	1.758	−0.457	−1.302	−0.500
			50	0.824	3.275	2.191	−1.587	−2.388	−2.609
	20	50	10	1.356	2.999	3.227	0.043	−0.520	−1.290
			50	2.490	3.036	2.907	−0.896	−0.642	−1.901
		250	10	0.255	3.227	4.534	−1.637	−1.157	−1.398
			50	0.748	2.839	3.436	−0.613	−0.092	−0.855

TABLE 27

RMS Localization Error as a Function of Burst Duration, Duty Cycle, and Source Azimuth for Experiment II

Burst duration (msec)	Duty cycle (percent)	Source azimuth (degrees) 20	60	100	260	300	340	Means
50	10	4.166	5.865	5.835	4.471	4.790	4.084	4.868
	50	4.479	5.284	5.253	4.787	4.828	3.483	4.686
250	10	3.786	5.821	6.071	4.881	4.561	3.640	4.793
	50	4.717	4.889	5.114	4.795	4.683	4.344	4.757
	Means	4.287	5.465	5.568	4.733	4.715	3.888	

TABLE 28

Linear Component of the Change in RMS Localization Error with Source Azimuth as a Function of Duty Cycle and Burst Duration

Burst duration (msec)	Duty cycle (percent)	
	10	50
50	(2 Hz)[a] -2.055	(10 Hz) -2.079
250	(0.4 Hz) -3.526	(2 Hz) -0.848

[a]Numbers in parentheses are the repetition rate for each signal, e.g., for a burst duration of 50 msec and 10 percent duty cycle, the bursts occurred two times per second.

linear component of the change in RMS over azimuth, as a function of Burst duration and Duty cycle. The negative scores indicate that RMS error was greater at ±100 degrees azimuth than at ±20 degrees. The numbers in parentheses above each data point are the rate of repetition of the bursts for each combination of Duty cycle by Burst duration. The low repetition rate (0.4 Hz) may have been responsible for the greater azimuth effect in the 250 msec burst, 10 percent duty condition. Subjects reported that they kept their head oriented toward 0 degrees azimuth until they heard the first burst on a given trial. Since localization accuracy decreased with azimuth relative to the head, a first burst at an extreme azimuth provided less information than a first burst at a small absolute azimuth. If the subject responded before a second burst occurred, localization should be less accurate at extreme azimuths. The fact that mean response time to ±100 degrees azimuth averaged only 0.8 sec greater than to ±20 degrees azimuth suggests that subjects often responded during the "off" part of the duty cycle (which lasted 2.25 sec in the 250 burst, 10 percent duty condition).

The only other effect for RMS error which approached significance was the three-way interaction of Rise time by Burst duration by Duty cycle ($F(1,3) = 5.92$, $p < .10$). Table 29 shows that, as expected, a shorter rise time produces greater localization accuracy (smaller RMS error) for three of four combinations of Duty cycle and Burst duration. The shorter rise time was most beneficial in the 250 msec burst, 10 percent duty, which had the lowest burst repetition rate. This effect is consistent with the general conclusion from lateralization studies that a short rise time is more beneficial when ongoing interaural time and amplitude differences are less available (due to the low repetition rate in this case). It is not clear why the shorter rise time failed to produce a beneficial effect in the 50 msec burst, 10 percent duty condition. The main effect of rise time was in the anticipated direction, but was not statistically reliable.

As mentioned earlier, it was of interest to study the trade-off (in terms of localization accuracy) between rise time and burst duration separately for high and low frequency stimuli. The three way interaction of Noise-type by Rise time by Burst duration for RMS error did not approach significance ($F(1,3) = 2.35$, $p = .2228$). Table 30 shows the RMS error data for these three variables.

TABLE 29

RMS Localization Error as a Function of Rise Time, Burst Duration and Duty Cycle

Duty cycle (percent)	Burst duration (msec)	Burst repetition rate (Hz)	Rise time	
			1	20
10	50	2	5.003	4.734
	250	0.4	4.491	5.095
50	50	10	4.512	4.859
	250	2	4.606	4.908
		Means:	4.653	4.899

TABLE 30

RMS Localization Error as a Function of Rise Time, Noise Type, and Burst Duration

Noise type	Burst duration (msec)	Rise time (msec) 1	20	Means
1 kHz low-pass	50	4.621	4.968	4.794
	250	4.787	4.853	4.820
2 kHz high-pass	50	4.894	4.625	4.760
	250	4.311	5.150	4.730

It was also of interest to determine the effect of repetition rate on localization accuracy for high and low frequency stimuli. There was no significant effect of repetition rate (Burst duration x Duty cycle) overall ($F(1,3) = 3.83$, $p = .1453$). Furthermore, there was no evidence of any differential effect of repetition rate for high versus low frequency stimuli, i.e., no significant interaction of Noise type by Burst duration by Duty cycle ($F(1,3) = 0.43$, $p = .5599$). Table 31 shows the RMS error data for these variables.

The ANOVA for mean response time produced two significant effects and one marginally significant effect. The linear trend with absolute source azimuth was significant. ($F(1,3) = 10.26$, $p < .05$). Table 32 shows that response times were about 0.8 sec faster on the average to sources at ± 20 degrees than at ± 100 degrees azimuth. Given that response times were generally long (the overall mean was 3.5 sec), the difference of 0.8 sec to localize in the front versus the rearward azimuths seems relatively small. The minimum response time was 0.714 sec, the maximum was 7.889 sec, and the standard deviation was 1.753 sec.

Another significant difference in mean response time was for Noise type, as shown in Table 33. Response times were, on the average, 0.6 sec faster for 1 kHz low-pass noise bursts than for 2 kHz high-pass noise bursts ($F(1,3) = 10.43$, $p < .05$). It is notable that there was a difference in the same direction, although smaller, for mean response times in Experiment I. In that experiment, the time required to localize a continuous, 1 kHz low-pass noise was about 0.2 sec less than that for a continuous, 2 kHz high-pass noise.

A third effect for mean response time in Experiment II approached significance. This was the interaction of Noise type by Duty cycle ($F(1,3) = 7.65$, $p < .10$). As can be seen in Table 33, mean response time decreased as duty cycle increased for the low frequency noise bursts, but increased with duty cycle for the high frequency noise bursts.

5.4 Discussion and Conclusions

In both Experiments I and II, there was relatively little difference in localization accuracy as a function of stimulus characteristics in the time and frequency domains. Tables 34 and 35 summarize the major results of Experiments I and II, respectively. On the basis of previous findings, it

TABLE 31

RMS Localization Error
as a Function of Noise Type,
Burst Duration, and Duty Cycle

Noise type	Burst duration (msec)	Duty cycle (percent) 10	50
1 kHz low-pass	50	(2 Hz)[a] 4.908	(10 Hz) 4.680
	250	(0.4 Hz) 4.958	(2 Hz) 4.682
2 kHz high-pass	50	(2 Hz) 4.829	(10 Hz) 4.691
	250	(0.4 Hz) 4.628	(2 Hz) 4.832

[a]Numbers in parentheses are the burst repetition rates.

TABLE 32

Mean Response Time as a Function of Source Azimuth

Source azimuth (degrees)					
20	60	100	260	300	340
3.035	3.386	3.721	4.062	3.748	3.145

TABLE 33

Mean Response Time as a Function of Noise Type and Duty Cycle

Noise type	Duty cycle (percent) 10	50	Means
1 kHz low-pass	3.326	3.094	3.210
2 kHz high-pass	3.684	3.961	3.822
Means	3.505	3.527	

TABLE 34

Summary of Major Results of Experiment I

- RMS Localization Error
  - White noise produced slightly, but significantly, less error than did 2 kHz high-pass noise
  - Localization tended to be slightly more accurate on the second day of practice than on the first day
  - Subjects localized sources on the right side more accurately than sources on the left side on the first day, but there was no difference on the second day
- Mean Response Time
  - Subjects responded significantly faster on the second day of practice than on the first day

TABLE 35

Summary of Major Results of Experiment II

- RMS Localization Error
  - There was a small, but significant, linear increase with source azimuth
  - The 1 msec rise time produced less error than did the 20 msec rise time for signals with the lowest repetition rate (0.4 Hz)
  - The increase in error with azimuth tended to be greater for stimuli with low repetition rates than for stimuli with higher repetition rates
- Mean Response Time
  - There was a small, but significant, linear increase with source azimuth
  - The 1 kHz low-pass noise stimuli produced slightly faster localization than did the 2 kHz high-pass noise stimuli

would be expected that such characteristics would exert a greater influence on localisation accuracy (see Section 2.0 for a review). However, almost without exception, the listener's head was restrained in the previous investigations. The relative absence of effects of stimulus characteristics in the present experiments suggests that head movement provides localisation cues which obviate the need for multiple cues provided by certain desirable spectral and time-domain characteristics (see Section 2.0 for a review of such characteristics).

Freedman and Fisher (1968) reported one of the few previous studies which relate the availability of spectral cues to localisation accuracy when the head is free to move. They studied localisation accuracy with and without head movement and with the listener's pinnae open or occluded. When the head was fixed, occlusion of the pinnae reduced localisation accuracy. When head movement was allowed, localisation was as accurate with occluded pinnae as with open pinnae. Thus head movement appears to greatly reduce the impact of stimulus characteristics on localisation accuracy, both with simulated cues and in normal unaided localisation.

Although the overall effects of stimulus characteristics on localisation performance were small, some trends and interactions were observed which have significance for the design of directional auditory displays. In Experiment II, the overall effects of the burst duration, duty cycle, and rise time of temporally intermittent stimuli on localisation accuracy and response time were not statistically significant. However, there was a trend for shorter rise times to produce greater localization accuracy, particularly for low repetition-rate stimuli. There was also a differential effect of repetition rate as a function of source azimuth. The decrease in localisation accuracy with absolute azimuth was greatest for very slow repetition rates, probably because the listener did not wait for a second burst before responding after turning his head in the general direction of the sound, as explained in Section 5.3.3.

Varying stimulus characteristics also had relatively little impact on the time required to localize a sound. There were no significant overall effects of high or low-pass filtering continuous noise stimuli, and no overall effects of rise time, burst duration, or duty cycle for intermittent stimuli. However, for intermittent noise stimuli, low-pass filtering did produce faster response times than did high-pass filtering. Furthermore, the advantage (in terms of shorter response times) for low-pass stimuli was accentuated at high duty cycles (slow repetition rates).

There was relatively little effect of absolute source azimuth on localisation accuracy and response time in Experiments I and II. For continuous noise stimuli, there was no effect of source angular distance from the frontal sagittal plane. For intermittent noise stimuli, there was a small but significant effect of absolute azimuth, localisation error increasing about 1.0 degree RMS for sources at ±20 degrees azimuth versus sources at ±100 degrees azimuth. The decrease in localisation accuracy with azimuth was greatest (about 1.75 degrees RMS) for very low repetition rate stimuli. There was also a small but significant increase in response time with absolute azimuth for intermittent noise stimuli. Mean response time increased about 0.8 sec for sources at ±20 versus ±100 degrees azimuth.

The fact that absolute azimuth has only small effects on localisation accuracy can be attributed to head movements. In the psychological fidelity experiment, the head fixed condition produced an average difference in localization error of 16.6 degrees RMS for the frontal region (+5 to -5 degrees) versus the lateral region (±130 to ±135 degrees). This difference is much greater than that obtained when the head was free to move (2.7 degrees). On the other hand, head movement does not seem to be accountable for the increase in response time with absolute azimuth. The fixed-head condition of the psychological fidelity experiment produced an average increase in mean response time of only about 0.7 sec for frontal versus lateral source positions, versus 1.0 sec for the moving head condition.

6.0 LONG-TERM RESEARCH AND DEVELOPMENT PLAN

This section includes a long-term plan for developing an electronic system which could convey accurate directional information by way of acoustic signals to listeners wearing headphones. The system will have the capability to impress directional qualities on incoming signals and messages in real time, and to alter the directional qualities of the sound in real time with head movement. The simulated auditory localization (SAL) cues delivered via headphones should enable a pilot to localize sound with accuracy comparable to that in free-field listening conditions.

6.1 Objectives

The overall objectives of the proposed program are to establish technical requirements, design, build, and test an electronic system for impressing a directional quality on audio signals and messages in the cockpit. The system is called a real-time directional synthesizer (RTDS). In order to build such a system, the following specific objectives must be accomplished:

1. Determine the minimum rate at which the system must sample head position and correct the apparent direction of the sound.

2. Determine the required spatial resolution of the system.

3. Determine the required bandwidth and dynamic range of the part of the system that synthesizes directional signals.

4. Measure and digitize transforms that relate free-field sound from given directions to the sound received at the ear canal entrance. This must be done at the minimum spatial resolution.

5. Fit the directional transforms, in either the time or frequency domains, with digital filter models.

6. Implement the models on a computer as non-real time digital filters and use the filters off-line to produce directional messages and sounds for testing.

7. Test the psychological fidelity of alternative filter designs and select the best filter models for further research.

8. Review previous research to determine individual differences in head-related transfer functions that are related to superior localization performance.

9. Modify the digital filter models to incorporate features associated with superior localization performance.

10. Evaluate localization performance with the modified filters as a function of background noise level with the listener's head fixed.

11. Evaluate the psychological fidelity of the modified filter in producing sound movement relative to the head.

12. Specify the engineering requirements for a real-time directional synthesizer (RTDS) and prepare a plan for development of a prototype.

13. Design the RTDS hardware and software.

14. Fabricate and bench-test the RTDS prototype.

15. Test the psychological fidelity of the RTDS prototype.

16. Document the RTDS hardware and software and recommend any needed refinements to the design.

The proposed work builds upon the previous work described in this report, and will use the SAL research facility described in Section 3.0.

6.2 Approach

In this section, alternative design approaches for an electronic system for producing directional audio signals in the cockpit are examined. The capabilities and characteristics that such a system must have include the following:

1. It must provide directional information in a manner which is not workload intensive. For the human, turning the head and eyes to visually acquire an object that has produced a sound is a natural, almost "automatic" response (cf., Posner, 1978; Shiffrin & Schneider, 1977).

2. It must accurately reproduce the acoustic cues available in free-field listening for binaural headphone presentation. A reasonable goal is to enable the pilot wearing earphones to localise sounds and messages with accuracy comparable to that with the unaided ear in a natural setting not characterised by high reverberation or noise.

3. The system, in conjunction with a system which measures head rotation and translation, should be able to modify the directional cues in real time with sufficient speed to compensate for head movement, aircraft motion, and motion of the object, if any, that the sound is signalling.

4. The system should minimise front-back reversals in localization judgements by accurately representing the acoustic cues associated with the external ears (pinnae) and head movement.

5. The engineering requirements of the system, in terms of information storage and access time, must be within the bounds of existing or near-term technology, and be such that the system is suitable for airborne applications.

6. It is highly desirable that the system be able to impress a directional quality on incoming messages, including voice communications, in real time. That is, the system should not be limited to reproducing prestored sounds and messages.

Two extreme approaches in terms of sophistication for building the required system are the following:

a. Off-line recording for later playback of a large number of sounds and messages from a large number of directions relative to a model of the human head, torso, and ears.

b. Real-time synthesis of directional cues based on modeled transforms for a large number of directions relative to the head.

The first approach has the advantage that it is technologically simpler. However, it is limited to reproducing prestored sounds and messages. It also has major problems in terms of the speed at which stored directional sounds must be accessed and in terms of the amount of storage required.

As the listener, in this case a pilot, turns the head, the directional sound delivered to the headset must be changed if the source of the sound is to appear to remain stationary, or move in a manner which is not slaved to head motion. A human can easily move the head at a rate of 90 degrees/sec. In order for the sound to appear to change smoothly, i.e., not jump back and forth, its directional qualities must be changed before the listener's head rotates through an angle equal to the minimum audible angle that the human can discriminate. The minimum audible angle for sounds located in front of the listener is about 1 degree. Hence, the directional sounds must be accessed at a rate greater than 90 Hz for sounds in this region. Of course, the direction in which the head will be moved cannot be predicted in advance. The system must therefore be able to access all the prerecorded directions for a given sound with the same speed, i.e., the system must have a "random" access capability.

The only storage medium which has the required random access capability and speed is solid-state random-access memory (RAM). An estimate of the amount of RAM required for this type of system is shown in Figure 29. The estimate of the number of directions from which each sound must be recorded is based on the auditory acuity measured in previous research (see Section 2.0 for a review). The required RAM for this approach, 96,000 megabytes, is not feasible for airborne applications with current or near-term technology. Even if the message length and number of directions were restricted, the system would still suffer the disadvantage that it cannot impress a directional quality on incoming, real-time messages.

- NUMBER OF BYTES OF RAM REQUIRED

  $= N \bullet M \bullet L \bullet S \bullet B$

WHERE N = NUMBER OF MESSAGES/SOUNDS

M = NUMBER OF DIRECTIONS (BASED ON SPATIAL RESOLUTION OF AUDITORY SYSTEM DURING HEAD MOTION)

L = AVERAGE MESSAGE LENGTH

S = SAMPLING RATE (AT LEAST 2 • BANDWIDTH)

B = NUMBER OF BYTES PER RECORD NEEDED

$$= \frac{\text{DYNAMIC RANGE}}{8 \text{ BITS/BYTE} \bullet 6 \text{ dB/BIT}}$$

$$\underset{N}{50} \bullet \overbrace{\underset{\text{Az}}{200} \bullet \underset{\text{EL}}{40}}^{M} \bullet \underset{\text{(SEC)}}{\overset{L}{3}} \bullet \underset{\text{(Hz)}}{\overset{S}{40{,}000}} \bullet \underset{\text{(BYTES)}}{\overset{B}{2}} = 96{,}000 \text{ MEGABYTES}$$

Figure 29. RAM requirements for off-line pre-recording approach.

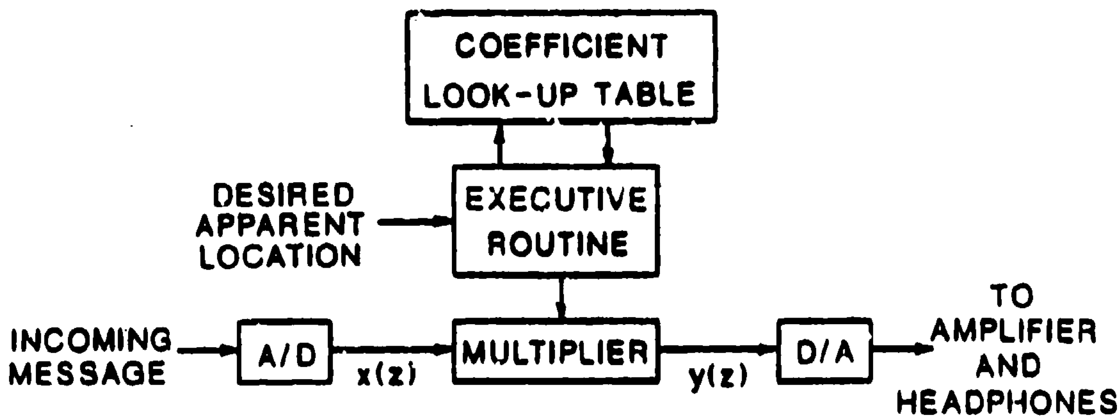
The second approach, real-time synthesis of directional cues, could be implemented by first modeling, off-line, the directional transformation that a sound undergoes from the free-field to the ear canal entrance for each of a large number of directions. Incoming messages could then be digitized (if they're not already in digital form) and digitally filtered using the modeled directional transform for the desired direction of arrival. Rather than storing directional messages, this approach would require storage of only a small set of coefficients for each direction.

The real-time operation of this approach is outlined in Figure 30. The system would perform the following functions in real time: (1) digitize the incoming message, (2) input the actual or desired direction of arrival of the sound relative to the aircraft, (3) compute the direction of the sound relative to the pilot's head, based on head position data from a head-tracking system, (4) look up the appropriate filter coefficients in RAM, (5) digitally filter the digitized, incoming message in either the host computer or in an outboard digital processor, and (6) convert the resulting waveform to analog form, amplify and low-pass filter it as necessary, and output to the pilot's headphones.

This approach has three major advantages relative to the first approach. First, much less RAM is required. Figure 31 shows the estimated amount of RAM needed. Second, access time to waveforms (in this case stored as model coefficients) is not a problem. Third, the system is not limited to pre-recorded sounds.

The approach also has two disadvantages relative to the first approach. The first is that a large number of directional transforms must be modeled. This is not a serious problem, since it need be done only once, off-line, for all future operations of the system. The second disadvantage is that the approach may require a host computer with high computational speed for the digital filtering operation. This could be overcome by using a set of special-purpose outboard microprocessors which perform the required multiplications in parallel.

In terms of both capabilities and developmental feasibility, the real-time synthesis of directional cues seem to be the preferable approach. The proposed research and development therefore addresses this approach.



$$Y(z) = H(z)\ X(z)$$

$$H(z) = \frac{a_0 + a_1 z^{-1} \cdots\cdots + a_N z^{-N}}{b_0 + b_1 z^{-1} \cdots\cdots + b_m z^{-m}}$$

Figure 30. Conceptual diagram of a real-time directional synthesizer showing the digital filtering operation.

- NUMBER OF BYTES OF RAM REQUIRED =

$$\left[(M \bullet K) + R\right] \bullet B$$

WHERE M = NUMBER OF DIRECTIONS

K = NUMBER OF COEFFICIENTS REQUIRED TO MODEL TRANSFER FUNCTIONS

B = NUMBER OF BYTES/COEFFICIENT OR PREVIOUS INPUT/OUTPUT VALUE

R = NUMBER OF PREVIOUS VALUES OF INPUT AND OUTPUT

$$\left[\left(\overbrace{\underset{Az}{200}\quad\underset{El}{40}}^{M}\quad\overset{K}{20}\right) + \overset{R}{20}\right] \cdot \overset{B}{2} = 320 \text{ KBytes}$$

Figure 31. RAM requirements for real-time synthesis approach.

6.3 Major Technical Issues

6.3.1 System Spatial Resolution

The real-time directional synthesizer must have sufficient resolution to make a sound appear to come from a single, stationary point in space, even when the listener moves his or her head. As the head rotates, the directional transform applied to the signal must be altered to represent the new direction of the sound source relative to the head. If the difference between the new and old directions of the sound source exceeds the minimum audible angle (MAA), the sound will appear to initially move in the direction opposite head movement and then jump in the direction of head movement. The MAA is the smallest angular distance which allows the listener to discriminate between two successive, stationary sounds (cf., Mills, 1958). If the angular distance between transforms is smaller than the MAA, then as the head moves, the listener should not detect an audible transition as the transform is changed.

Perrott and Musicant (1977, 1981) studied the MAA with moving sound sources. Listeners were asked to adjust the time of onset for a moving sound source, such that the onset occurred directly in front of them. For source velocities from 45 to 120 degrees/sec, the MAA measured in this way was comparable in magnitude to the MAA measured by Mills (1958). Thus it appears that the MAA is a valid measure of auditory acuity for moving as well as static sound sources.

Although many previous investigations have measured auditory acuity, the methods, stimuli, and source directions studied have differed greatly. As a result, our knowledge of human auditory acuity is incomplete. Many of the more recent studies have used dichotic presentation and, for theoretical reasons, have used unnatural stimuli which would never be encountered in free-field conditions. The results of these studies are therefore not useful for determining the proper angular interval at which transform must be measured for the real time synthesizer.

The most comprehensive study of free-field acuity was conducted by Mills (1958). He studied the MAA for pure tones as a function of azimuth in the horizontal plane of the interaural axis. Gardner and Gardner (1973) measured localization accuracy for connected speech stimuli in the median sagittal plane and for other vertical planes at various angles to the median plane. Wettschureck (1973) measured the MAA for white noise at various elevations in the median plane. Harris (1972) replicated Mills' results for the 0° azimuth region, but found somewhat smaller MAA's for ±30° azimuth. Searle, Braida, Davis and Colburn (1976) developed a model of auditory localization based on data from studies of horizontal and vertical plane localization reported over a ten-year period. Their model predicts the standard deviation of localization error (which is equivalent to the MAA) for the frontal horizontal plane and the median plane as a function of the span of the loudspeaker array and the acoustic cues available. Their model is actually a sophisticated summary of the available MAA data. If MAA data were available for other angular regions, the model could be extended to those regions. It could serve as a tool for estimating the MAA between measured directions, and perhaps reduce the number of directions for which data must be collected in this program.

Although previous research will provide a useful guide, systematic measurements of the MAA are needed at closely spaced intervals for virtually all combinations of azimuth and elevation. From previous research on auditory localization (see Section 2.0 for a comprehensive review) it appears that localization accuracy is as great or greater for white noise than for any other auditory stimulus. White noise is therefore the logical choice as a standard stimulus for the quantification of the MAA.

Task 1 of Phase II consists of experimental research to fill the gaps in the available data on the MAA. The MAA will be measured for white noise at increments of azimuth and elevation over the entire spherical space surrounding the listener's head. The MAA should govern the required resolution of the synthesizer when the movement of the sound source relative to the head is slow. When the movement of the head and/or sound is more rapid, the auditory apparent movement illusion may provide a better means of producing the perception of smooth movement of the sound relative to the head. The exploitation of the apparent movement phenomenon is discussed in the next section.

6.3.2 System Response Time

In order to be useful in the cockpit, the real-time synthesizer must be integrated with a system which: (a) measures head translation and rotation at a high sampling rate, and (b) rapidly computes the appropriate sound direction relative to the head. This information must be output to the real-time synthesizer, which will produce a sound with the appropriate directional qualities. The response time of such a system, from the pilot's execution of a head movement to the production of the sound at the new apparent direction, must be short enough to make the sound appear stationary during head movement and to simulate the smooth movement of sound sources about the head.

The simplest algorithm for achieving perceptually smooth movement of the sound relative to the head would be to successively transition between transforms which characterize directions differing by less than one MAA. This approach puts very heavy demands on the response time of the system. A human listener can easily turn his or her head at a rate of 90°/sec. For sound sources located in the anterior median plane, where most objects of interest to a pilot would be, the MAA is as small as 1° (Mills, 1958). In order to keep up with head movement, the system would need to transition among transforms and produce appropriate sounds at a rate of 90 Hz.

An alternative approach to minimizing system response time is to synthesize only a reduced set of localization cues during rapid head movements. Freedman & Fisher (1968) have suggested that pinnae cues do not contribute to localization accuracy during head movements, but do help direct the initial orienting response (head movement) to a sound. They measured localization accuracy in the horizontal plane with listeners using their own pinnae, with their pinnae occluded, and with artificial pinnae, with or without head movement allowed. With no head movement, their own or artificial pinnae increased localization accuracy considerably relative to the no-pinnae condition. However, when head movement was allowed, there was no significant difference between the pinnae and no-pinnae conditions.

This result suggests that the modulation of binaural cues during head movements is so powerful that it swamps out any effect of monaural cues. If this is the case, then it should be possible to simulate only the changing interaural time difference (ITD) and interaural amplitude difference (IAD) during rapid head movements with no loss of localization accuracy. It would be relatively simple to vary the ITD of the signal reaching the pilot's headphones as a function of head orientation in real time. Since the IAD increases as a function of frequency, it would be necessary to high-pass filter the signal in a direction-dependent manner. Such a filter would be much simpler, however, than a filter which simulates all the localization cues. Also, it might not be necessary to alter the characteristics of the IAD high-pass filter at MAA-size increments as the sound moves relative to the head. The filter could be changed as a continuous function of the angle of the sound relative to the interaural axis. Reducing the cues simulated should greatly reduce the amount of processing needed during rapid head movements and therefore reduce the required system speed.

An interesting question is whether the "reduced cue" approach could be used to simulate movement of a sound about a stationary head. The acoustic information received by the ear should be the same regardless of whether the head moves relative to the sound or vice-versa. It is possible, but not likely, that proprioceptive feedback from the neck muscles changes the way in which auditory information is processed centrally, i.e., somehow promotes the use of pinnae cues. Even if this were the case, the reduced cue approach would still be useful. In the present application, the maximum angular rate of head movement is much greater than the angular rate at which an object would be expected to move about the pilot and aircraft. If necessary, the two types of movement could be simulated by different methods. When the head is moving rapidly one could use the reduced cue approach; when the head is stationary or slowly moving, one could successively transition among transforms associated with directions differing by less than one MAA.

A third possible strategy is to exploit the apparent movement (beta) phenomenon. Taking advantage of the apparent movement phenomenon would make it possible to synthesize rapid movements of a sound relative to the head by successively activating transforms which are separated by more than one MAA.

When two spatially separated stimuli are presented in succession, the observer may perceive not two stimuli, but a single stimulus which moves from the position of the stimulus which appeared first to the position of the second stimulus. In order to obtain apparent motion, the interval between the onsets of the two stimuli, the distance between the stimuli, and the intensity of the stimuli must be in the proper proportion. The phenomenon also depends on the duration of the stimuli.

Wertheimer (1912) and Korte (1915) reported extensive studies of apparent movement in the visual modality (cited in Murch, 1973). According to Korte, as the distance between the stimuli is increased, the interstimulus onset interval (ISOI) must be increased proportionately in order to maintain the apparent movement. When the observer perceives one object moving smoothly from the position of the first stimulus to the second, the phenonenon is called "beta" movement. As the ISOI is lengthened, a point is reached at which the movement no longer appears smooth, but broken or jerky. The broken movement is called the "phi" phenomenon. As the ISOI is lengthened further, the observer eventually perceives two spatially separated stimuli appearing successively. Visual apparent movement plays an important role in virtually all types of visual simulation, including such mundane applications as motion pictures.

The apparent movement phenomenon also occurs in the auditory modality (Burtt, 1917; Mathiesen, 1931; Briggs & Perrott, 1972; Geldard, 1984). Each of these studies confirmed the general relationships described by Korte among the ISOI, the distance separating the stimuli, the intensity of the stimuli, and the stimulus duration. However, the optimal ISOI for the auditory modality (25 to 30 msec for stimuli 10 to 60 centimeters apart) appears to be shorter than the optimal ISOI for visual stimuli (about 75 to 125 msec for stimuli a few centimeters apart).

Among the studies employing acoustic stimuli, the ratio of the distance between the stimuli to the ISOI differed greatly. This ratio is important in the present application because it is related to the rate at which the sound moves relative to the head, which must be precisely controlled. The differences among the studies are probably attributable to differences in the methods and stimuli used. In the earlier studies the equipment was quite crude by modern standards.

Further research is needed to determine how to take advantage of the apparent movement phenomenon in the present application. Questions of interest include the effect of the type of stimulus, the angluar distance over which apparent movement can be achieved, the rates of movement which are possible, how the foregoing variables differ as a function of the angular region in which the movement occurs (e.g., in front vs. in the back of the head), and individual differences in the optimal conditions for apparent auditory movement.

6.4 Proposed Research

The proposed research is organized into three phases. The overall goal of the first phase is to determine the feasibility of the digital-filter approach to developing a real-time directional synthesizer. This will be accomplished by providing preliminary answers to five major technical issues. The issues include: (1) the minimum system response time required, (2) the spatial resolution required, (3) the required bandwidth of the synthesizer, (4) the required dynamic range of the synthesizer, and (5) individual differences in head-related transfer functions (HRTF's) which are correlated with superior localization performance, and therefore might be used to enhance performance for all listeners. It is anticipated that the third and fourth issues can be answered satisfactorily by reviewing previous research findings. The remaining issues will require additional experimental research. Phase I will include the reviews of previous research for all five issues and preliminary experimental research on the first issue.

The overall goal of Phase II is to determine the engineering requirements for the real-time directional synthesizer. The first task will examine the spatial resolution required in such a synthesizer in order to maintain psychological fidelity during head movements and preserve head-movement related cues to the location of the sound source. Also in this phase, the time-domain waveforms containing directional information will be measured and digitized. This data will then be used to evaluate alternative digital filter models. Next, the best-fitting digital filter models will be implemented on a computer so that they can be evaluated in nonreal time. A small number of messages will be digitized and convolved with the alternative nonreal time digital filters to produce messages which have a directional quality. These directional messages will then be used in a variety of experimental tests. One set of tests will examine the psychological fidelity of alternative filter models in terms of localization accuracy and speed. The results of these tests will be used to select one or two filter models for further experimental testing.

A second set of experiments will examine the utility of emphasizing selected features of the HRTF's in terms of localization accuracy and speed. Individual differences in HRTF's that are associated with superior localization performance will be emphasized by altering the digital filter. These modified HRTF's will be tested with human listeners to determine which features should be retained in the filter model. The results of these two sets of experiments relate directly to the engineering requirements for maximizing the static fidelity of the synthesizer. That is, for maximizing localization performance when both the head and sound source are fixed.

A third set of experiments will examine the psychological fidelity of the system when the sound source moves with respect to the head (i.e., in conditions of head movement relative to a stationary sound, or sound movement about a stationary head, or a combination of the two).

At the end of Phase II, the major design requirements for the real-time directional synthesizer will have been determined. The last major task in Phase II will be to summarize these requirements, including recommendations for any further research needed, and to develop alternative design approaches for building a real-time directional synthesizer.

In the third phase of the proposed program, a design will be developed for the real-time directional synthesizer. A prototype will be built and bench-tested. This phase will also involve the measurement of a large number of directional waveforms. A series of experimental tests will be conducted in Phase III to evaluate the psychological fidelity of the prototype synthesizer and to optimize the synthesizer's performance. Recommendations will be made for any needed refinements to the synthesizer design and for any necessary further testing of the synthesizer.

7.0 APPENDIX

ANNOTATED BIBLIOGRAPHY

Articles were selected for annotation based upon their relationship to the goals of the literature review. In this regard, the majority of articles included concentrated on auditory localization models and literature reviews, horizontal and median-plane localization cues, auditory distance cues, effects of head movement and vision on localization, perception of auditory motion, and bisensory interaction. A small sample of articles involved noise effects, alternate environments, auditory displays, stereophonic listening principles, and acoustic manikin specifications. A topical index which cross-references the annotations according to the above categories is included at the end of this appendix.

1. Alekseenko, N. Yu. (1983). Role of movements of different types in spatial hearing acuity. Human Physiology, 8, 240-243.

The acuity of spatial hearing (i.e., the minimal audible angle) was evaluted with respect to spatial location (0° vs. 50° azimuth), type of response (verbal vs. pointing), and head orientation (facing forward or facing to the side). For verbal responses, performance was better at 0° than at 50° azimuth. Pointing improved performance at 50°. Head orientation had no effect.

2. Andreassi, J. L., & Greco, J. R. (1975). Effects of bisensory stimulation on reaction time and the evoked cortical potential. Physiological Psychology, 3, 189-194.

This study compared bisensory stimulation to unisensory (visual or auditory) stimulation, using both reaction time (RT) and evoked potential (EP) as dependent variables. Two methods were used to determine the temporal offset between the two stimuli in the bisensory condition. The first, ΔR, set the offset equal to the difference between the visual and auditory RTs. The second, ΔN2, set the offset equal to the difference between the visual and auditory N2 latency component of the EP measure. The results indicated that visual RT was faster than auditory RT, and bisensory stimulation was faster than the visual response. The EP measure resulted in shorter latencies in audition, and the amplitude of bisensory stimulation was higher than in unimodal stimulation. It was concluded that the increased amplitude of the bisensory stimulation condition resulted in facilitation of response time.

3. Batteau. D. W. (1967). The role of the pinna in human localization. Proceeding of the Royal Society, London, Series B, 165, (pp. 158-180). London, England.

A theoretical analysis on the role of the pinna in sound localization was presented. The pinna introduces quantitatively different time delays on incoming signals dependent upon the location of the source. These transformations, it is proposed, are used by the auditory system to recreate, via an inverse transform of the time delays, the location of the sound. The mathematical process of encoding these time delays was shown to be physiologically possible, based upon a fairly simple neural net of excitation and inhibition.

4. Batteau. D. W., Plante, R. L., Spencer, R. H., & Lyle, W. E. (1963). Localization of sound: Part 3. A new theory of human audition. (Report No. TP3109, Part 3). China Lake, CA: U. S. Naval Ordinance Test Station.

A theoretical analysis of sound localization, based upon the role of the pinna as a means of introducing time delays on incident auditory stimulation, was presented. Time delay transformations enable the auditory system, using information theory metrics, to perform monaural autocorrelations and binaural crosscorrelations to encode the signal source. The advantages of a time domain theory over the classical frequency/intensity theory include mechanisms to account for monaural localization, selective attention, localization without head movements, and masking phenomenon. The report emphasized the importance of high fidelity equipment and the acoustic properties of materials in recording time delay and developing artificial pinnae. Empirical evidence found time delay differences with changes in sound source azimuth and elevation, and applied theory successfully by testing underwater localization ability.

5. Batteau, D. W., Plante, R. L., Spencer, R. H., & Lyle, W. E. (1965). Localization of sound: Part 5. Auditory perception. (Report No. TP3109, Part 5). China Lake, CA: U. S. Naval Ordinance Test Station.

The basilar membrane is implicated, and mathematically modeled, as the structure responsible for the encoding of time delays imposed through pinna reflections on auditory stimuli. This refinement in delay theory resulted in the investigation of human speech recognition factors, and selective attention. The development of effective electrostatic headphones for use in accurately transmitting localization information was presented, and performance specifications given.

6. Biguer, B., Jeannerod, M., & Prablac, C. (1982). The coordination of eye, head, and arm movements during reaching at a single visual target. Experimental Brain Research, 46, 301-304.

The latency of eye, head, and arm movements directed at the same visual target were measured in five human subjects. Latency of activation of the corresponding neck and arm muscles was also measured. Although the overt movements occured sequentially (eye movement, head movement, then arm movement), EMG discharges were synchronous with respect to eye movement onset. The authors suggest that the relative synchrony of neural commands for eye, head, and arm movements, in producing the patterned sequence of overt movements, may have implications for eye-hand coordination (i.e., subjects made larger pointing errors when not allowed to move their head and eyes toward the target).

7. Blauert, J. (1969/1970). Sound localization in the median plane. Acoustica, 22, 205-213.

The theory of timbre differences was applied to auditory localization in the median plane. The theory proposed that linear distortions in the frequency spectra, caused by pinna, head, and ear canal reflections, changed dependent upon the angle of incidence of the sound source. In the first experiment, listeners reported whether a source appeared above, in front, or behind them. The results showed that localization was a function of frequency. Low frequencies (< 630 Hz) and frequencies between 2500 and 6300 Hz were perceived as coming from the front, frequencies between 800 and 2000 Hz and between 8000 and 12000 Hz were perceived as behind the subject, and frequencies between 6300 and 8000 Hz were perceived as above the listeners. A second experiment measured the frequency

components of sound at the eardrum, and confirmed that the frequency changes were indeed correlated with the perceived location. In the third experiment, recordings from the ear canal were played back to the listeners and were perceived as in experiment 1. These studies were taken as support for the directional band theory of median plane localization.

8. Blauert, J. (1971). Localization and the law of the first wavefront in the median plane. <u>The Journal of the Acoustical Society of America</u>, <u>50</u>, 466-470.

This study investigated and evaluated the precedence effect according to the directional band theory of localization. The precedence effect is the suppression of another sound when it follows a primary sound between 600 μsec to 50 msec. Below 600 μsec the second sound is summed with the first to create phantom sound sources; above 50 msec the second sound is perceived as an echo of the first. Using the method of constant stimuli, Blauert found that judgements of stimulus location (above, front or behind) did not, as expected, follow the predictions of the precedence effect when the secondary source followed the primary sound by less than 500 μsec. Blauert interpreted these results as confirming the relevance of the directional band theory of localization. Measurement of the frequency spectrum, with various delay intervals, revealed that linear distortions in the spectrum were consistant with the predictions of a combfilter transformation.

9. Blauert, J. (1981). Lateralization of jittered tones. <u>The Journal of the Acoustical Society of America</u>, <u>70</u>, 694-698.

This study reviewed the hypothesis that auditory lateralization was based on interaural time differences (ITDs) for complex high frequency signals. It was proposed that lateralization could be accomplished, even when the signal does not contain any components below 1600 Hz and the signal envelopes are flat, via a peripheral auditory filter which converts FM to AM signals. The output of this transformation would produce time-varying envelopes in previously flat signals. This hypothesis was evaluated using jittered tones. In contrast to Nordmark (1976), who found ITDs as small as 2 microseconds in jittered tones, these results (173 microseconds) did not differ from previous ITD thresholds found with non-jittered FM tones. It was concluded that the results did not contradict the FM to AM transformation hypothesis, but that no new auditory mechanism was necessary to explain localization with jittered versus non-jittered tones.

10. Blauert, J. (1982). Binaural localization. In O. J. Pedersen and T. Poulsen (Eds.), <u>Binaural Effects in Normal and Impaired Hearing</u> (Scandinavian Audiological Supplement 15, pp. 7-20).

This review evaluated the relationship between physical characteristics of sound, as transformed by the external ear, and the various psychophysical aspects of auditory localization cues. The purpose was to develop a conceptual model of signal processing

consistent with the above physical-psychophysical relationship. Blauert concluded that the use of dichotic listening tasks was inadequate in elucidating binaural interactions in sound localization. It was further suggested that more use be made of manikins with external ears in the identification of binaural cues.

11. Burkhard, M. D., & Sachs, R. M. (1975). Anthropometric manikin for acoustic research. Journal of the Acoustical Society of America, 58, 214-222.

The development of the KEMAR anthropometrically designed manikin was reviewed. Anthropometric data and KEMAR physical specifications (head size, orientation, and pinna dimensions) were emphasized. The acoustical characteristics of the KEMAR eardrum simulator are also identified and validated against human auditory data. The KEMAR manikin adequately simulates acoustic diffraction and responses in the 50th percentile adult population.

12. Butler, R. A. (1969a). Monaural and binaural localization of noise bursts vertically in the median sagittal plane. The Journal of Auditory Research, 3, 230-235.

Localization in the vertical plane, due to the lack of interaural phase or intensity differences, is often viewed as being difficult and inaccurate. However, if high frequency, complex sound sources are utilized, performance can be quite accurate. Butler investigated the role of possible binaural cues, based upon pinna transformations, using stimuli which varied in frequency composition. In all stimulus conditions binaural performance was more accurate than monaural performance. In addition, broad band noise and high pass noise produced better performance than low band pass noise. The results support the possible use of pinna transform cues in the vertical plane, and the necessity of high frequency, complex tones to make use of these cues.

13. Butler, R. A. (1969b). On the relative usefulness of monaural and binaural cues in locating sound in space. Psychonomic Science, 17, 245-246.

It was hypothesized that since binaural cues are available in the horizontal plane, but not the vertical plane, performance differences would reflect the increased number of cues available in a binaural task. In comparing horizontal to vertical localization performance, it was concluded that the binaural cues in the horizontal plane were the causal factor which improved localization accuracy above that found in the vertical plane.

14. Butler, R. A. (1970). The effect of hearing impairment on locating sound in the vertical plane. International Audiology, 9, 117-126.

This article examined localization accuracy in the vertical plane, using three groups which differed as to their hearing deficiencies. Subjects with a bilateral hearing impairment for frequencies in the region of 8000 Hz were compared to subjects with a unilateral hearing impairment and a normal hearing group. The bilateral hearing impaired group was completely unable to localize in the vertical plane, but was proficient in the horizontal plane. The unilateral impairment group performed significantly better, but not as well as the normal hearing group. The results support the notion that the availability of high frequency information is important in localizing in the vertical, but not the horizontal, plane. The difference between the monaural impairment and normal groups suggested that binaural cues were used in vertical plane localization.

15. Butler, R. A., & Belendiuk, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. The Journal of the Acoustical Society of America, 61, 1264-1269.

Three experiments investigated the role of spectral cues in the localization of sound in the vertical plane. Noise bursts were recorded via microphone implants in the external ear under free field conditions. Localization accuracy when the stimulus was presented through headphones was identical to free field conditions. Spectral analysis, summed over subjects, revealed a notch in the frequency response which was dependent upon frequency. Comparisons among subjects showed that the frequency spectra recorded in the ears of highly accurate subjects consisted of a regular notch which decreased with frequency, while the poor localizer's notch was not as orderly. This notch in the sound spectrum is apparently related to vertical plane localization accuracy.

16. Butler, R. A., & Flannery, R. (1980). The spatial attributes of stimulus frequency and their role in monaural localization of sound in the horizontal plane. Perception and Psychophysics, 28, 449-457.

The effects of frequency composition in monaural localization were investigated. It was found that as the center frequency of a 1 kHz wide band noise increased, within limits, the apparent location of the sound moved from 0 to 270 degrees azimuth. However, at certain critical frequencies the apparent location would return to 0° and then progress again towards 270° as the center frequency increased. This pattern of responses was interpreted as separate spatial referents, which were dependant upon frequency composition. Individual differences in terms of the ranges of the spatial reference maps (SRMs) were evident. Each subject displayed either two or three separate SRMs. SRM 1 tended to range between 4 and 8 kHz, SRM 2 ranged from 8 to 12 kHz, and SRM 3 ranged from 13 kHz and above. In a second experiment the location of the center frequency was varied such that for each subject the frequency range fell either within a single SRM or between SRMs. Accuracy was better in the between-SRM condition.

17. Butler, R. A., & Helwig, B. A. (1983). The spatial attributes of stimulus frequency in the median sagittal plane and their role in sound localization. American Journal of Otolaryngology, 4, 165-173.

This study investigated the role of spatial referents in the median plane. The apparent location of a sound source was dependent, not

upon its actual location, but rather on the center frequencies of 1.0 kHz wide noise bands. As the center frequency increased from 4 to 12 kHz, the apparent location of the sound shifted from in front to the back of the subject. At 13 kHz the sound again shifted to the front. Differences in spatial reference maps was attributed to differences in the pinna amplification function.

18. Chasson, L. R., McFarland, T. P., & Aldrich, T. B. (1971). Auditory effects on spatial orientation (Report No. 71-10). Colorado Springs, CO: United States Air Force Academy. (AD 737351)

The purpose of this research was to investigate the use of beat changes as a cue to maintain spatial orientation. The task required subjects to adjust the position of a rod via a joystick until the rod was parallel with a reference frame. Visual cues present or absent and auditory beat cues present or absent were fully crossed between subject variables. The dependent variable was absolute angle error. The results indicated that the presence of both auditory and visual cues produced the most accurate performance. The use of visual cues alone was only marginally more accurate than auditory cues alone.

19. Coleman, P. D. (1962). Failure to localize the source distance of an unfamiliar sound. The Journal of the Acoustical Society of America, 34, 345-346.

The role of experience in localizing the distance of a sound source was investigated. The results indicated that, on the first trial, distance estimates were unrelated to the actual source distance. However, with repeated exposures, the ability to accurately judge the distance of the source increased. It was concluded that distance estimation, in this case, was based upon the relative comparison of sounds on successive trials, rather than on absolute distance cues.

20. Coleman, P. D. (1963). An analysis of cues to auditory depth perception in free space. Psychological Bulletin, 60, 302-315.

Cues for auditory distance perception, and their supporting evidence, were reviewed. The attenuation of the amplitude of a sound wave with increases in distance was confirmed as a monaural cue to distance estimation when familiar sounds were employed. The frequency spectra of a sound source was also implicated as a monaural cue to distance estimation. At far distances (> 4 feet) the wave front is approximately planar and provides distance cues based due to the differential attenuation of high frequencies, relative to lower frequencies. This exponential loss is dependent upon humidity, temperature, terrain features, and inhomogeneities in the atmosphere. Binaural cues, such as intensity and phase differences between the two ears, were also investigated as possible distance cues. The ratio of sound pressure at the two ears (i.e., the binaural intensity ratio, or BIR) changes with distance as a function of azimuth. The change in distance necessary to produce noticeable changes in BIR values increases as either azimuth departs from 90 degrees, frequency increases, or distance increases. It was concluded that BIRs could not be a useful cue beyond 15 feet even with optimal azimuth, frequencies, and environmental factors. It was further concluded that distance estimation with binaural cues may permit absolute judgements of distance, while monaural cues tend to require experience with the sound source to estimate distance.

21. Coleman, P. D. (1968). Dual role of frequency spectrum in determination of auditory distance. The Journal of the Acoustical Society of America, 44, 631-632.

It was suggested that the frequency spectrum of a sound may play a dual role in auditory depth perception. At distances greater than a few feet, relatively greater high-frequency content may signify a closer sound source due to the greater attenuation of high frequencies relative to low frequencies by passage through air. At these distances the wavefront may be considered planer, and the differential attenuation follows the exponential loss law. At closer distances, however, the sound field must be considered spherical, and it is the low-frequency content of the spectrum that is more prominently affected by changes in distance. As distance is increased in this range, the low-frequency content is reduced more than the high-frequency content. Therefore, given a stationary sound source in the near field, a relative increase in the high frequencies may cause the source to appear more distant, whereas in the far field a relative increase in the high frequencies may cause the source to appear closer. The article cites experimental evidence in support of this dual role for the frequency spectrum in perception of auditory distance.

22. Dirks, D. D., & Gilman, S. (1979). Exploring azimuth effects with an anthropometric manikin. Journal of the Acoustical Society of America, 66, 696-701.

The KEMAR manikin's eardrum response to both pure tones and pink noise was measured, and comparisons were made with Shaw's (1974) data on the human eardrum response. The KEMAR manikin's eardrum response was comparable to Shaw's data, except for higher frequencies (≃ 8,000 Hz), and where the test ear was shadowed by the head (≃ -60 to -120 degrees).

23. Easton, R. D. (1983). The effect of head movements on visual and auditory dominance. Perception, 12, 63-70.

Two experiments were performed to evaluate the effects of visual discordance (created by looking through a prism) on a location task. In each experiment the subjects located a small audio speaker unimodally or bimodally. Half the subjects in each experiment were unaware they were viewing through a prism; the other subjects were informed that a prism was used and observed a demonstration of its refracting properties. The second experiment differed from the first in that head movement was allowed. The results of the first experiment indicated that knowledge of the visual discordance greatly reduced the bias toward the visual modality in bimodal trials; that is, the subjects tended to rely on the auditory cues more if they were aware their vision was distorted. In the second

experiment, head movements did not affect the precision of visual localization, but did reduce the visual bias in bimodal trials (from the levels observed in the first experiment) for no-knowledge subjects.

24. Feddersen, W. E., Sandel, T. T., Teas, D. L., & Jeffress, L. A. (1957). Localization of high-frequency tones. The Journal of the Acoustical Society of America, 29, 988-991.

The role of interaural time and amplitude differences (ITD and IAD respectively) cues in auditory lateralization was investigated. Physical measurements of ITD and IAD over various frequencies and azimuthal positions were made. The subjects were presented with a noise stimulus of preset ITD, and a pure tone whose IAD could be adjusted. The task was to match the apparent location of the noise and pure tone. The results indicated considerable disparity between the matched pure tone and the actual physical measurements recorded from the corresponding azimuthal position. The disparity was quite pronounced at low frequencies and decreased considerably for higher frequencies.

25. Firestone, F. A. (1930). The phase difference and amplitude ratio at the ears due to a source of pure tone. The Journal of the Acoustical Society of America, 2, 260-270.

The amplitude and phase of a pure tone were measured at the two ears of a manikin for various azimuthal and distance positions, and the amplitude ratio and phase difference were calculated. In all cases, as distance decreased the amplitude ratio decreased and the phase difference increased. Azimuthal position effects were dependent upon distance and frequency. In general, the largest phase difference for signals below 1000 Hz was found at 90 degrees, and the minimum amplitude ratio was greatest at ≈ 90 degrees.

26. Flannery, R., & Butler, R. A. (1981). Spectral cues provided by the pinna for monaural localization in the horizontal plane. Perception and Psychophysics, 24, 438-444.

The apparent location of 1 kHz wide band noise burst was found to be dependent on the center frequency of the band. As frequency increased (4 to 9 kHz) the apparent source location migrated from 0 to 90 degrees. At the higher frequencies the source appeared, for some subjects, to revert back to 0° and proceed again towards 90°. These patterns of apparent location were labeled spatial reference maps (SRMs). Measurement of pinna amplification showed a positive relationship between the apparent location and the amplification function.

27. Forbes, T. W. (1946). Auditory signals for instrument flying. Journal of the Aeronautical Sciences, May, 255-258.

In an aircraft simulator, with both naive and experienced pilots, auditory signals were successfully combined to supplement visual information. The three-in-one signal indicated heading variations

by a directional sweep of the sound, bank information by pitch, and airspeed by rate of occurance. The results indicated that the three-in-one signal can be successfully used, with practice, to maintain simulated flight. An automatic system for verbally announcing various instrument readings was described. Five principles for the successful implementation of auditory signals in the cockpit were presented and discussed.

28. Freedman, S. J., & Fisher, H. G. (1968). The role of the pinna in auditory localization. In S. J. Freedman (Ed.), The Neurophysiology of Spatially Oriented Behavior. (pp. 135-152). Homewood, IL: Dorsey Press.

A series of experiments was undertaken to examine the role of the pinna in auditory localization. Subjects restrained from head movements were able to accurately localize only if their own pinnae or artifical pinnae were present. Subjects with one ear occluded were able to localize accurately. Even when the interaural time difference cues were confounded by lengthening the interaural axis, subjects could still localize as long as artifical pinnae were used. These findings suggest the pinna plays a major role in auditory localization and cast doubt on the signficance of localization experiments which eliminate the role of the pinna through the use of earphones.

29. Freedman, S. J., & Pfaff, D. W. (1962). The effect of dichotic noise on auditory localization. The Journal of Auditory Research, 2, 305-310.

This study investigated the effects of noise and motility on a dichotic time discrimination task. The time discrimination task consisted of obtaining difference thresholds both before and after exposure to noise and motility conditions. Three motility conditions were employed in a within-subjects design. These were (1) ambulatory - the subject walked through a busy corridor in 5 minute periods while listening to dichotic noise stimuli, (2) recumbent - the subject layed in bed without gross body movements, and (3) passive movement - the subject was wheeled through the busy corridor as in the ambulatory condition. It was predicted, based upon visual displacement studies, that performance in dichotic time discrimination would decrease as noise exposure and active movement increased. These results were supported.

30. Freides, D. (1974). Human information processing and sensory modality: Cross-modal functions, information complexity, memory and deficit. Psychological Bulletin, 81, 284-310.

The theoretical and empirical relationships between sensory modality and information processing were reviewed. It was suggested that performance differences between modalities are dependent upon the informational complexity of the stimulus. Performance tends to be equal, regardless of modality, if the stimuli are relatively simple. More complex stimuli tend to show a performance advantage for a particular modality.

31. Gardner, M. B. (1969a). Distance estimation of 0° or apparent 0°-oriented speech signals in anechoic space. The Journal of the Acoustical Society of America, 45, 47-53.

The ability of observers to estimate the distance of speech signals was investigated in a series of experiments. The apparent distance of the source was varied by manipulating either the intensity of the source or the actual distance from the source to the listener. Both recorded and live voices were used as signals. Using live voices which varied in intensity (i.e., whisper, quiet speech, conversational speech, or shouting), observers were quite accurate at judging the distance of the sound, although whispered speech tended to be underestimated and shouted speech tended to produce overestimation. Recorded voices, in all conditions, tended to produce poor distance estimates.

32. Gardner, M. B. (1969b). Image fusion, broadening, and displacement in sound localization. The Journal of the Acoustical Society of America, 46, 339-349.

The various methodologies for producing "phantom" sound locations in stereophonic (two speaker) systems are reviewed. Techniques for producing stereophonic effects from monophonic signals are discussed.

33. Gardner, M. B. (1973). Some monaural and binaural facets of median plane localization. The Journal of the Acoustical Society of America, 54, 1489-1495.

This article examined the localization of noise-band signals in the anterior sector of the median plane. Both subjective observations and objective measurements (using a manikin head) were obtained in an anechoic chamber. Optimum localization accuracy was obtained in binaural conditions with the cavities of both pinnae entirely unoccluded. In monaural conditions the apparent location of the sound tended to shift away from the median plane; hence, binaural reception seems to be important in maintaining the proper azimuthal position for median plane sources. Occlusion of the cavities in the pinna (leaving the ear canal open) was found to degrade localization, suggesting that the pinna cavities provide important monaural cues. Measurements taken at the manikin eardrum revealed that amplitude differences, as a function of elevation angle, are too small to account for more than a small part of the localization accuracy obtained in this sector of the median plane. Differences in frequency response characteristics were found to be more substantial, and thus a possible source of cues.

34. Gardner, M. B., & Gardner, R. S. (1973). Problem of localization in the median plane: effect of pinnae cavity occlusion. The Journal of the Acoustical Society of America, 53, 400-408.

This study examined the effect of pinna cavity occlusion on localization in the anterior and posterior sectors of the median plane. Results indicated that (1) localization accuracy decreases as the degree of pinna cavity occlusion increases, (2) accuracy is

better for broad-band noise signals than narrow-band signals for all degrees of occlusion, (3) accuracy increases as the frequency of the noise band increases, and (4) greater accuracy is possible in the anterior than the posterior sector of the median plane. It was concluded that differences in pinna transforms, as a function of elevation angle, are an important source of cues for localization in the median plane.

35. Garner, W. R. (1949). Auditory signals. In A survey report on human factors in undersea warfare (pp. 201-217). Washington, DC: National Research Council.

This review discussed the feasibility of using auditory signals to reduce visual workload in a submarine scenario. The various types of auditory information were discussed, and basic psychological attributes employed in encoding the auditory information were emphasized. The advantages and disadvantages of auditory signalling were identified.

36. Gatehouse, R. W. (1982). Summary: New directions. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 267-270). Groton, CT: Amphora Press.

The summary remarks identified several key areas of research needed in auditory localization. Generally, the need to study ecologically valid, and dynamic (rather than static) localization situations was stressed. In part, research needs in distance location, moving sound and moving listener effects, duration of stimulus effects, and masking effects were identified.

37. Gopher, D. (1973). Eye-movement patterns in selective listening tasks of focused attention. Perception and Psychophysics, 14, 259-264.

This article examined the role of eye movements in dichotic listening tasks. During the presentation of auditory messages the spontaneous occurrances of eye movements were reduced, and the eye movements and fixations which did occur tended to be in the direction of the ear receiving the message. When the dichotic switching task was changed to a monaural presentation mode, the eye movements during message presentation tended, unlike before, to remain centrally located. It was concluded that the elimination of dichotic competition reduced the orienting response towards the signal. In an effort to increase listening demands, monaural presentation, with increased message frequency, was included. The results indicated that increasing the difficulty of the task did not increase the eye movement orienting response.

38. Gotoh, T. (1982). Can the acoustic head related transfer function explain every phenomenon in sound localization. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 244-248). Groton, CT: Amphora Press.

An application for psychoacoustically derived head-related transfer functions to improve stereophonic reproduction was presented. Both diotic and dichotic transform information was necessary to produce accurate sound localization from a stereophonic system.

39. Grantham, D. W., & Wightman, F. L. (1978). Detectability of varying interaural temporal differences. The Journal of the Acoustical Society of America, 63, 511-523.

This study investigated the effects of interaural time differences (ITDs) in detecting intracranial movement. Two parameters of the stimuli were varied. These were the modulation frequency (fm) that is, the rate of movement, and the peak sinusoidal difference in ITDs (ΔT) that is, the extent of movement. The results, using a 2 forced choice procedure, indicated that the magnitude of ΔT necessary for movement detection, increased as fm increased. The encoding of temporal cues was characterized as a relatively slow, low-pass filter.

40. Greene, D. C. (1968). Comments on "Perception of the range of a sound source of unknown strength." The Journal of the Acoustical Society of America, 44, 634.

This paper, using Hirsh's (1968) equation to determine the distance of a sound source, showed that the relative error, under optimal conditions, would produce 45% uncertainty in range estimate. It was concluded that Hirsh's equations were unreliable as a cue for distance localization.

41. Hebrank, J., & Wright, D. (1974a). Are two ears necessary for localization of sound sources on the median plane? The Journal of the Acoustical Society of America, 56, 935-938.

Localization in the median plane has often been attributed to pinna transformations of the incoming source, due to the lack of interaural intensity and time differences. However, binaural localization has been found to be superior to monaural localization, indicating that the binaural comparison of the wavefronts at each ear is a necessary cue for localization. Experiment 1 compared binaural and monaural performance with either familiar or unfamiliar sounds. Binaural performance, as expected, was superior to monaural, and familiar sounds were more accurately localized than unfamiliar sounds. Experiment 2 evaluated the change in performance in monaural localization with training and feedback. The training resulted in monaural localization accuracy improving to previously measured binaural performance. The necessity of two ears for median plane localization was rejected.

42. Hebrank, J., & Wright, D (1974b). Spectral cues used in the localization of sound sources on the median plane. The Journal of the Acoustical Society of America, 56, 1829-1834.

Three experiments investigated the role of frequency spectra on sound localization in the median plane. The experiments identified the features encoded by the external ears. The results, as expected, differed according to the frequency spectra presented. Frontal localizations were perceived with one octave notch whose lower cutoff frequency was between 4 and 8 kHz. Localizations above the listener occurred for sound sources of 1/4 octave bands which peaked between 7 and 9 kHz. Locations behind the listener were perceived with frequency bands which peaked between 10 and 12 kHz. In general, the cues for median plane localization reside between 4 and 12 kHz; however the specific relationship between notch frequency and perceived auditory elevation was not clear.

43. Herschenson, M. (1962). Reaction time as a measure of intersensory facilitation. Journal of Experimental Psychology, 63, 289-293.

Reaction times to sounds, lights, and sound-light combinations, were measured. In the light and sound combination condition the light always preceded the sound, and the degree of onset asynchrony was varied. The results indicated that the light and sound combination facilitated reaction time, and that the greatest facilitation occured when the asynchrony was equal to the difference in reaction times between the independent visual and auditory conditions.

44. Hirsh, H. R. (1968). Perception of the range of a sound source of unknown strength. The Journal of the Acoustical Society of America, 43, 373-374.

It was mathematically demonstrated that the distance of a sound, regardless of its initial amplitude, could be calculated by interaural time and amplitude differences between the two ears. The relationship between interaural differences and the distance of a source was not supported by any psychophysical data. Therefore, the utility of this calculation cue is unknown. A hypothetical psychophysical experiment was proposed to explore this possibility.

45. Holt, R. E., & Thurlow, W. R. (1969). Subject orientation and judgement of distance of a sound source. The Journal of the Acoustical Society of America, 46, 1584-1585.

Noise bursts, which controlled for loudness cues, were presented at various distances from an observer. Subjects whose right ear faced towards the sound source were able to accurately rank-order the distance of the stimuli. However, subjects who directly faced the sound source were unable to discriminate between distances. These results have implications for interaural time and amplitude difference cues.

46. Jones, B., & Kabanoff, B. (1975). Eye movements in auditory space perception. Perception & Psychophysics, 17, 241-245.

The effects of eye movements and their correspondance to auditory information were investigated using a signal detection paradigm. Experiment 1 evaluated the effects of localization with eye movements fixed or free, and found that eye movements improved localization. Experiment 2 introduced congruent or discrepent visual information to assess the effect of eye movements. The results indicated a significant difference between congruent and discrepent information, but the effect of eye movements was not significant. Therefore, it was suggested that it is not eye movements, per se, that affect localization performance, but rather the congruence between eye movements and location. Experiment 3

substituted visual information with verbal instructions to move their eyes one direction or the other, in order to assess the effects of eye movements in the absence of visual cues. The results again support the notion that eye movements affect the localization of auditory signals even in the absence of visual information.

47. Kock, W. E. (1950). Binaural localization and masking. The Journal of the Acoustical Society of America, 22, 801-804.

This study investigated the role of interaural time differences (ITDs) in the discrimination of binaurally presented speech signals from noise. In general, a reduction in ITD increased the difference threshold for identifying speech. It was concluded that differences in arrival cues allow for sound localization, and improve discriminability of signals from noise.

48. Kuhn, G. F. (1977). Model for the interaural time differences in the azimuthal plane. The Journal of the Acoustical Society of America, 62, 157-167.

Interaural time and amplitude differences (ITD and IAD, respectively) were measured in a KEMAR manikin at various azimuthal positions. Comparisons between measured and theoretical ITD values were made, and were concluded to be reasonably similar. The data support that ITD is independent of frequency both below 500 Hz and above 3000 Hz. ITDs were minimal between 1400-1600 Hz for azimuthal positions between 15 and 60 degrees. These results support past findings, which concluded that localization was poor between 1000-2000 Hz, and that the localization cue of IAD was used above 1400 Hz.

49. Kuhn, G. F. (1982). Towards a model for sound localization. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Applications (pp. 51-64). Groton, CT: Amphora Press.

A mathematical model of sound localization cues, based on interaural time differences (ITD) for low frequency information, and interaural amplitude differences (IAD) for high frequency information, was presented. The model was supported by changes in ITD and IAD for various hearing aid placements.

50. Lambert, R. M. (1974). Dynamic theory of sound-source localization. The Journal of the Acoustical Society of America, 56, 165-171.

The effects of head movement on the localization of sound source azimuth and range was mathematically modeled. In binaural situations, interaural time differences (ITD) will systematically change as head position is changed. The difference between ITDs on two successive samples of the stimulus, and knowledge of the degree of head movement from sample 1 to sample 2, will generate mathematical solutions for determining azimuth. In terms of binaural range perception, mathematical solutions to distance based upon interaural amplitude differences (IAD) sampled between successive head movements were generated. Attention was paid to the limitations imposed by human interaural time and amplitude

difference thresholds, and the parameters (range and angle) over which these cues could operate. In terms of theory it is unclear if observers calculate these changes or use learned relationships to determine azimuth and range. In addition, a two factor theory of localization was proposed, in which low-frequency information is utilized if head movements are available, and high-frequency information is utilized when head movements are constrained.

51. Levy, E. T., & Butler, R. A. (1978). Stimulus factors which influence the perceived externalization of sound presented through headphones. The Journal of Auditory Research, 18, 41-50.

In two experiments, the stimulus factors involved in the externalization of sound presented through headphones was examined. Experiment 1 varied the interaural time differences (ITD), interaural spectral differences (ISD), and interaural amplitude differences (IAD) available in the stimulus. Three stimulus tapes were created: Tape 1 contained ITD, ISD, and IAD cues, tape 2 contained ITD and ISD cues, and tape 3 contained only ITD cues. In each condition frequency composition of the noise bursts were varied from .5 to 6 kHz. The results indicated a significant effect of frequency composition, with high frequency sounds being judged as closer to the head than low frequency sounds. No effect of interaural cue content was found. The second experiment decomposed ITD cues into interaural arrival differences (IARD) and interaural ongoing differences (IOND). One tape contained both IARD and IOND cues (i.e., tape 3 from experiment 1) and the second eliminated IARD cues from tape 3. The results showed a significant effect of frequency as in experiment 1, and a significant difference between frequency composition cues. The IOND cues alone condition was perceived as being closer to the head than the IOND and IARD condition. It was concluded that the perception of externalized sound can occur even when many of the natural cues are eliminated.

52. Loudsbury, B. F., & Butler, R. A. (1979). Estimation of distances of recorded sounds presented through headphones. Scandinavian Audiology, 8, 145-149.

Two experiments investigated the role of frequency composition, monaural vs. binaural presentation, and azimuthal position in an auditory distance estimation task. Subjects received noise bursts (270 degrees azimuth) recorded at the eardrum of other individuals. The stimuli varied as to the amount of head shadow available from distances of 2 to 10 feet. Stimuli were presented, in a within subjects design, both monaurally and binaurally for 1 kHz low-pass noise bursts, unfiltered noise bursts, and 4 kHz high-pass noise bursts. Subjects were given one reference tone at 5 feet as a comparison stimulus. The second experiment was identical to the first, except that azimuthal position was varied (360, 330, 300, and 270 degrees). Nine out of 48 subjects in the two experiments tended to reverse their judgements of distance. These subjects, termed inverters, were analyzed separately from the non-inverters. The results showed that non-inverters accurately estimated the distance of the sound source, especially at higher frequencies and at 330 degrees azimuth. Monaural versus binaural presentation had little effect upon performance. It was concluded that the intensity ratio of direct to reflected sound can, when loudness cues are unavailable, be used as a relative cue for distance estimation.

53. Loveless, N. E., Brebner, J., & Hamilton, P. Bisensory presentation of information. Psychological Bulletin, 73, 161-199.

The literature concerning bisensory presentation of information was reviewed and interpreted according to signal detection theory. In general, it was assumed that the addition of redundant information would, due to probability summation, reduce detection thresholds. Four experiments, using a signal detection paradigm, were conducted to investigate unimodal and bimodal performance as the information redundancy between the signals is varied. In general, the results indicated that redundancy between the signals is important; however, the effect differs widely between subjects.

54. Mastroianni, G. R. (1982). The influence of eye movements and illumination on auditory localization. Perception and Psychophysics, 31, 581-584.

This study critically evaluated Shelton and Searle's (1980) conclusions on visual facilitation of auditory localization. An experiment was designed to evaluate the frame of reference hypothesis and memory stabilization hypothesis. The crossed independant variables in a repeated measures design were illumination (light or dark) and eye movements (permitted or not permitted). An ANOVA on mean error scores indicated that only eye movements in a lighted environment improved performance. In general, the results tended to lend more support to the frame of reference position.

55. Matin, L. (1982a). Visual and auditory localization: Normal and abnormal relations. In G. T. Chisum & P. E. Morway (Eds.), Research Program Review: Aircraft Physiology (Report No. NADC-82232-60). Warminster, PA: Naval Air Development Center.

This article reviewed several studies on how the coordination of bisensory localization is accomplished and identified possible mechanisms. The studies, using auditory/visual matching tasks, used three populations of subjects: strabismic, bilateral asymmetric hearing deficits, and normal subjects. The article defined visual capture, visual-field suppression, and cancellation mechanisms for visual localization. Three theoretical cancellation mechanisms - inflow, outflow, and hybrid theory - were identified.

56. Matin, L., Picoult, E., Stevens, J. K., Edwards, M. W., Young, D., & MacArthur, R. (1982). Oculoparalytic Illusion: Visual-field dependent spatial mislocalizations by humans partially paralyzed with curare. Science, 216, 198-201.

The oculoparalytic illusion (which occurs under curare-induced paralysis) was discussed. The illusion occurs only in darkness, and can be described as a misperception of the location of a fixated light. Restoration of illumination restores the appearance of the light to its actual position, and apparently no amount of training influences this illusion. Evaluation of the cause of the illusion centered upon whether the angle of the eye in the head or the angle of the head and body with respect to gravity were involved. Systematic variations of both angles concluded that the former was responsible. Another study evaluted the effects of this illusion on auditory localization. In normal illumination, paralyzed observers were accurate in positioning a visual target to the median plane. However, in matching a sound source location to a light, large gaze-dependent errors occurred.

57. Matin, L., Stevens, J. K., & Picoult, E. (1983). Perceptual consequences of experimental extraocular muscle paralysis. In A. Hein & M. Jeannerod (Eds.), Spatially Oriented Behavior. New York, Springer-Verlag.

Complete data is provided on the causal relationship between the oculoparalytic illusion and the angle of the eye in the head. In terms of auditory-visual matches, curarized subjects in full illumination tended to make localization match errors related to gaze position. The errors were attributed to a suppression of the output of extraretinal eye position information.

58. McFadden, D., & Pasanen, E. G. (1975). Binaural beats at high frequencies. Science, 190, 394-396.

Evidence is presented which suggests that binaural beats (previously obtained only for low-frequency sinusoidal inputs) can be experienced at higher frequencies if complex waveforms are used. The phenomenon of binaural beats is characterized by the perception of a single beat created by the presentation of two separate tones, which differ in frequency, to different ears. In some ways the binaural beats found for high-frequency complex waveforms were similiar to beats perceived with low-frequency sinusoids. However, there were several ways in which they differed. High frequency beats tended to be fainter, to be perceived over a larger range of frequencies, and to alter pitch perception. It was concluded that the beating phenomenon is probably based upon the same mechanisms that account for low-frequency beats.

59. McFadden, D., & Pasanen, E. G. (1976). Lateralization at high frequencies based on interaural time differences. The Journal of the Acoustical Society of America, 59, 634-639.

A series of experiments investigated the effects of high frequency complex waveforms on auditory lateralization. A single-interval-forced-choice paradigm was employed. The task consisted of determining the side to which two signals, one of which was time delayed to either the right or left ear, were lateralized. In general, the results support the hypothesis that the auditory system, contrary to the classical duplex theory of localization, is sensitive to ongoing interaural time differences (ITDs) at high frequencies. Several parameters of the wavefront were investigated. These were center frequency, use of two-tone complex

signals, depth of modulation, and simultaneous presentation of complex wavefronts which differed in center frequency. It was concluded that the auditory system is sensitive to ITD at high frequencies when complex sound sources are employed, and that the analysis of ITD effects must differentiate between arrival differences, ongoing time differences, and envelope time differences.

60. McNulty, J. A. (1982). Underwater sound and human hearing. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Applications (pp. 250-266). Groton, CT: Amphora Press.

The ability of humans to localize sounds underwater was assumed to be relatively poor. It was felt that the change in medium would distort the classical localization cues such as head shadow and interaural time and intensity differences. In a series of seven experiments, the ability to localize sounds and use echo information to determine the distance of a sound source, while underwater, was shown. Accuracy was increased through training, practice, and under anechoic conditions (i.e. open ocean).

61. Mehrgardt, S., & Mellert, V. (1977). Transformation characteristics of the external human ear. The Journal of the Acoustical Society of America, 61, 1567-1576.

Transfer function (amplitude and phase) measurements were taken over a frequency range of 200 to 15000 Hz for both the horizontal and median planes. Measurements were made 2. mm inside the ear canal. The results agreed well with previous studies. At azimuths between 0 and -54° (away from the ear) an amplitude dip is present at 1 kHz, due to head diffraction. For positive azimuthal values, the amplitude increased relative to the free field. In the vertical plane, the only increase in amplitude is at 8 kHz, at 90°. This corresponds to the finding that noise bands in this region are perceived as from above. These results have implications for the transformations (phase and amplitude) the pinna performs on incoming signals.

62. Metcalfe, J., Glavanov, D., & Murdock, M. (1981). Spatial and temporal processing in the auditory and visual modalities. Memory & Cognition, 9, 351-354.

Three studies evaluated the effects of input modality and type of recall. Visual and auditory modalities, and spatial and temporal recall were fully crossed variables in a repeated measures design. In all three experiments spatial recall was superior to temporal recall in the visual modality, and temporal recall was superior to spatial recall in the visual modality. It was concluded that each modality is sensitive to different aspects of the stimulus array.

63. Mills, A. W. (1958). On the minimum audible angle. The Journal of the Acoustical Society of America, 30, 237-246.

This study measured the difference threshold for locating the direction of a source. Using tonal stimuli, ranging from 250 to

10,000 Hz, it was found that the minimum audible angle (MAA) was dependent upon both frequency and azimuthal position (measured between 0 and 90 degrees). The MAA is approximately 1 degree for a source at 0° azimuth and between 500-750 Hz. Increases in azimuth increased the MAA. The effects of frequency were dependent upon azimuth, but in general tended to be cyclic in nature. In the median plane the effects of frequency were similiar to that found in the horizontal plane. Interaural phase/time differences (ITD), for low frequencies, and interaural amplitude differences (IAD), for higher frequencies, were obtained and compared at various MAAs at different azimuthal positions. The results indicated that the ITD threshold, under optimal MAA conditions, was approximately 10 μsec. The IAD threshold, again under optimal MAA conditions, averaged .5 db.

64. Mills, A. W. (1963). Auditory perception of spatial relations. In Proceedings of the International Congress on Technology and Blindness, Vol. 2 (pp. 111-139). New York: American Foundation for the Blind.

This review identified various factors in auditory localization, and discussed neuronal models to account for them. A comprehensive review of the work of Feddersen, Sandel, Teas, and Jeffress (1957), and Sandel, Teas, Feddersen and Jeffress (1955) on the localization of pure tones revealed that the localization of low frequency tones (< 1500 Hz) was accomplished primarily through interaural time differences (ITDs). The poor localization of tones around 1500 Hz was attributed to phase ambiguities, the gradual breakdown of neuronal synchrony above 1000-1500 Hz, and the lack of ITD cues based on the waveform envelope for the sound waves used in the studies. This review also covered the work of Wallach (1940) on head movement effects. Echo localization and the precedence effect was also reviewed. Dallenbach's (1953) work on echo reflections was reviewed, as were the natural echoes produced by the pinna (Batteau, 1963, 1965). It was hypothesized that these pinna reflections could be peripherially encoded on the basiliar membrane if it was assumed that some, but not all, of the neurons fired each time a deflection was made on the basiliar membrane. The two sets of time varying neurons would then combine information into an autocorrelation matrix. The autocorrelation function would encode the pattern of excitation, similiar to models of pitch perception encoding, and deduce the source location.

65. Molino, J. (1973). Perceiving the range of a sound when direction is known. The Journal of the Acoustical Society of America, 53, 1301-1304.

Hirsh's (1968) derivation of the role of interaural time and amplitude differences in the perception of distance was questioned. In essence, the correlation between interaural time and amplitude differences (ITD and IAD, respectively) results in the distance estimate being contingent upon both the source amplitude and azimuth. In order to use Hirsh's equation relating IAD and ITD, one of the above parameters must be independently available to the observer. Hirsh's model was tested by using an array of speakers at various distances (3 to 48 feet) which were calibrated to control for the 1/R Loss Law. The results indicated, contrary to Holt and Thurlow's (1969) data, that observers with their right ear towards the speaker were unable to make reliable judgements based on ITD and IAD information. The discrepancy between experiments was attributed to procedural and measurement differences.

66. Morimoto, M., & Ando, Y. (1982). On the simulation of sound localization. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 85-98). Groton, CT: Amphora Press.

The simulation of sound localization using measured head related transfer functions (HRTF) of three subjects was reported. Subjects were tested in both the median and horizontal plane. The simulated localization, in the horizontal plane, produced results comparable to real sound sources when the subject's own HRTF was used. When another subject's HRTF was used, localization ability decreased. Simulated localization was more accurate in the horizontal plane than the median plane.

67. Moushegian, G., & Jeffress, L. A. (1959). Role of interaural time and intensity differences in the lateralization of low-frequency tones (Report No. DRL-A-144). Austin, TX: The University of Texas. (AD-A031-955).

This study investigated interaural time and amplitude differences (ITD and IAD, respectively) in low frequency (500 - 1000 Hz) tones. Using the method of adjustment, a subject was given a standard tone, and adjusted the ITD of a noise stimulus until it appeared to occupy the same intracranial position as the tone. The results indicated that as the stimulus intensity increased, the amount of ITD necessary to match the tone location decreased. This effect was discussed in terms of encoding latency and stimulus intensity. However, differences found in the time-intensity functions for different frequencies and subjects indicate that other processes are involved.

68. Mudd, S. A., & McCormick, E. J. (1960). The use of auditory cues in a visual search task. Journal of Applied Psychology, 44. 184-188.

This article investigated use of auditory signals to reduce visual search time. Three parameters of the auditory signal were combined into one-, two-, or three-dimensional cues used to identify a sector or groups of sectors in which a deviant dial setting was located. The results indicated that two- and three-dimensional cues significantly reduced search time over the unidimensional group and the no-cue control group. A significant sector-by-group interaction was found. Apparently the no-cue group tended to search the array via a left to right, top to bottom strategy, whereas the cued groups' strategy was dependent upon the auditory cue.

69. Mulligan, R. M., & Shaw, M. L. (1981). Attending to simple auditory and visual signals. Perception & Psychophysics, 30, 447-454.

This study investigated the changes in performance when attentional demands were switched between trials. One group was presented with either auditory or visual information constant within a block, while another group was presented with either auditory or visual information which changed from trial to trial. The effects of increased probability of error in the divided attention group, due to set size, were partialled out of the analysis. The results indicated no decrement in performance due to attentional switching demands. It was concluded that the difference between conditions was due to non-attentional (set size) effects.

70. Musicant, A. D., & Butler, R. A. (1984a). The influence of pinnae-based spectral cues on sound localization. The Journal of the Acoustical Society of America, 75, 1195-1200.

Two experiments were performed to investigate the role of the pinna in providing spectral cues for binaural localization. In experiment 1, noise bursts of four types (broadband, 4 kHz high pass, 4 kHz low pass, and 1 kHz low pass) were presented to subjects with either both pinnae occluded or both open. The results showed a significant decrease in localization error with open pinnae for the broadband and 4.0 kHz high pass noise stimuli. Experiment 2 investigated the role of the near or far pinna in front-rear localization reversals. High pass noise bursts were presented to subjects with either both pinnae occluded, both pinnae open, the right (far) pinna occluded, or the left (near) pinna occluded. In terms of front-rear reversals in localization, the results suggested that the near pinna is responsible for perception of the quadrant of the source. The authors concluded that a theory of binaural localization must incorporate pinnae-based cues to account for increased accuracy and encoding of front or rear quadrant location.

71. Musicant, A. D., & Butler, R. A. (1984b). The psychophysical basis of monaural localization. Hearing Research, 14, 185-190.

This study investigated monaural localization of wide band noise bursts, at various center frequencies, in the horizontal plane. The apparent location of the stimuli were not dependent upon actual position, but rather was based on frequency composition. As the center frequency increased, the apparent location of the stimulus shifted from 330 to 225 degrees. Further, it was found that the location judgments were related to the speaker location which would have resulted in the greatest pinna amplification of critical frequencies. This relationship, it was concluded, is based upon an internal spatial referent which associates the location of the stimulus to the specific pinna amplification of the frequency bands.

72. Nickerson, R. S. (1973). Intersensory facilitation of reaction time: Energy summation or preparation enhancement. <u>Psychological Review</u>, <u>80</u>, 489-509.

This article reviewed studies which showed that the addition of secondary information (usually auditory) would facilitate reaction time to the primary stimulus (usually visual). The addition of a second stimulus improves performance in responding to the primary stimulus even when it is irrelevant to the task. Nickerson evaluated these results in terms of whether the second stimulus increased physiological energy summation or was an alerting cue as to the primary stimulus. It was concluded that a preparation-enhancement hypothesis was adequate to explain the findings.

73. Nordlund, B. (1962). Physical factors in angular localization. <u>Acta Otolaryngology</u>, <u>54</u>, 75-93.

Time, phase, and amplitude differences between the two ears of a dummy head were measured. Azimuth and frequency of tonal stimuli were varied. Time and phase differences were found to be a direct function of azimuth between 0-60° and 120-180°. Amplitude, on the other hand, tended to be an irregular function of azimuth. Comparisons were made with the theoretically based predictions of Hartley and Fry (1921) and with the empirical results of Feddersen, et. al. (1955), Firestone (1930), Mills (1958), and Sivian and White (1933). The article is a comprehensive and consise reference for binaural time, phase and amplitude differences.

74. Oatman, L. C. (1975). <u>Simultaneous processing of bisensory information</u> Aberdeen, MD: U. S. Army Human Engineering Laboratory. (AD-A012 149)

The effects of bisensory processing of information were reviewed with respect to redundancy of information, vigilance, information rate, and detection difficulty. Bisensory facilitation was found to occur when the information was redundant, but only when the stimuli were difficult to detect, recognize, or discriminate. Performance improves in bisensory tasks only if the information is redundant. It was concluded that non-redundant information in a bisensory task could, in high workload situations, unnecessarily increase response demands.

75. Perrott, D. R. (1982). Studies in the perception of auditory motion. In R. W. Gatehouse (Ed.), <u>Localization of Sound: Theory and Application</u>, (pp. 169-193). Groton, CT: Amphora Press.

Studies on the perception of illusory auditory motion were reviewed. Using apparent movement in the visual system as a model, similiar auditory movement effects (e.g. autokinesis, induced motion, motion aftereffects, auditory blur, and velocity judgements) were identified. Auditory movement paradigms have produced perceived movement with either a stationary source and listener, a moving source with a stationary listener, a stationary source with a moving listener, or a moving source and listener. In dichotic listening, the perception of movement was found to be a direct

function of interstimulus onset interval (ISOI) and signal duration. The optimal ISOI for perception of continuous auditory motion increased as signal duration increased. The autokinetic effect occurs in both the vertical and horizontal plane, and its duration was found to be a decreasing function of increases in signal bandwidth. Motion aftereffect creates the perception of movement through the use of a stationary listener and source with a moving auditory background. After prolonged exposures, the auditory source is perceived as moving in the direction opposite the background. The paper also reviewed the perception of auditory velocity. In general, listeners were very consistent in their velocity judgements (power function slope =1.0).

76. Perrott, D. R., & Elfner, L. F. (1968). Monaural localization. The Journal of Auditory Research, 8, 185-193.

Monaural localization was compared in two conditions: one in which loudness differences between the two speakers were equated, and one in which they were not. Performance was also compared to a binaural condition. The results indicated that while binaural performance was nearly errorless, monaural performance with speakers matched for loudness was at chance, and that monaural performance with unequal loudness levels fell in between. A second experiment was conducted which provided training for the subjects, and added a second condition which reversed loudness (i.e., with the left ear occluded the left speaker's intensity was greater than the right speaker). The results supported the hypothesis that an intensity cue was being used. In the reversed loudness condition subjects consistantly selected the speaker with the greater intensity. No support was found for pinna based cues being utilized in either experiment. When loudness cues were held constant performance was at chance.

77. Perrott, D. R., & Musicant, A. D. (1981). Dynamic minimum audible angle: Binaural spatial acuity with moving sound sources. The Journal of Auditory Research, 21, 287-295.

The Minimum Audible Angle (MAA), originally developed by Mills (1958), was expanded to encompass moving sound source stimuli. The method of adjustment was used, whereby the subject could adjust the initiation of a signal until it was perceived to be at 0 degrees azimuth. Velocity was varied over four conditions: 45, 60, 120, and 240 degrees/sec. The results showed that dynamic MAA was comparable to the static MAA developed by Mills except at the highest velocity. It was concluded that the minimum angular difference threshold of moving sources is similar to the MAA derived under static conditions.

78. Pick, H. L., Warren, D. H., & Hay, J. L. (1969). Sensory conflict in judgements of spatial direction. Perception and Psychophysics, 6, 203-205.

Using an artifical discrepancy paradigm, the dominance of one modality over another was evaluated. The modalities investigated were vision, audition,and proprioception. The results indicated visual information biased audition and proprioceptive judgments, and proprioceptive information biased audition judgments. It was concluded that combinations of sense modalities do not behave as an integrated system; rather, in certain circumstances, some modalities influence responding more than others.

79. Platt, B. B., & Warren, D. H. (1972). Auditory localization: The importance of eye movements and a textured visual environment. Perception and Psychophysics, 12, 245-248.

Two experiments were conducted which investigated the role of eye movements, in either light or dark conditions, and eye-hand coordination in an auditory localization task. Experiment 1 showed that eye movements under lighted conditions were more accurate in locating an auditory signal than eye-fixation and no vision conditions. Experiment 2 investigated the role of eye movements and pointing in lighted, dark, and fixation conditions. Saccadic eye movements were more accurate in the lighted eye movement condition than the dark eye movement condition. In addition, the correspondance between eye and hand position was greater in the lighted eye movement condition. It was concluded that target-directed eye movements produced the most accurate localization due to the availability of a textured environment to supply information about eye position.

80. Pollack, I., & Rose, M. (1967). Effect of head movement on the localisation sounds in the equatorial plane. Perception and Psychophysics, 2, 591-596.

This study, in five experiments, investigated the conditions under which head movement improved localisation. In general, the conditions are (1) sustained sound sources, and (2) sound source locations at large angles outside the median plane. They conclude that head movement facilitates localisation by allowing the listener to take advantage of the higher-acuity region around 0° azimuth.

81. Rodgers, C. A. (1981). Multidimensional localization: An investigation of possible pinnae cues. Dissertation Abstracts International, 42, 1811B.

Measurements of head related transfer functions (HRTF), at 24 horizontal positions, were made for three subjects, and were compared in both the time and frequency domains. The HRTFs differed between subjects, but showed similiar trends. For two of the three subjects, the HRTFs tended to systematically vary as a function of source position. Possible cues produced by the pinna transformations are discussed.

82. Roffler, S. K., & Butler, R. A. (1968). Localization of tonal stimuli in the vertical plane. The Journal of the Acoustical Society of America, 43, 1260-1266.

A series of experiments investigated the role of frequency and body orientation in localization of tonal stimuli in the vertical plane. It was found that localization in the vertical plane was a function of frequency. Higher frequencies tended to be judged as

originating higher in the vertical plane; lower frequencies were judged as originiating lower in the plane. Several additional experiments investigated the role of vision and body orientation in making these judgements. Manipulations included changing subject orientation with regard to the vertical plane, decreasing visual angle through increases in distance from the plane, using congentially blind subjects, and using children without knowledge of high and low descriptions of pitch. In general, the phenomenon remained consistent.

83. Russell, G. (1976). The role of the pinna in monaural horizontal plane localization. Journal of Auditory Research, 16, 68-70.

Binaural localization was compared to monaural localization with and without pinna cues (open or occluded pinna). Performance was poorer in monaural conditions than binaural conditions. Monaural localization was worse with the pinna occluded than with the pinna open.

84. Sandel, T. T., Teas, D. C., Feddersen, W. E., & Jeffress, L. A. (1955). Localization of sound from single and paired sources. The Journal of the Acoustical Society of America, 27, 842-852.

The role of interaural time differences (ITDs) in auditory localization was investigated. Tones ranging from 500 to 5000 Hz, and a wide-band noise, were used as stimuli. The task consisted of localizing a sound, presented by either one or two speakers, by adjusting the location of the noise source until the noise and tone were perceived as occuring at the same location. Three experiments were performed. The first, using a single source, evaluated the accuracy of the tonal-noise matching task. Experiments 2 and 3 used two speakers, and differed as to whether the tonal stimuli were presented in phase (Experiment 2) or out of phase (Experiment 3). ITD measurements were used as the basis for predictions of localization. The results indicated that the tonal matching method produced accuracy above that found in pointing methods (i.e., intermodal matching), and that the predictions of location, based upon ITD vector analysis, were consistant with the data up to 1500 Hz. It was concluded that ITD was the major cue utilized in low frequency localization.

85. Searle, C. L. (1982). A model of auditory localization: Peripheral constraints. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Applications (pp. 42-50). Groton, CT: Amphora Press.

A block diagram model of the physical and neural processes involved in auditory localization was presented. The decision mechanism was hypothesized to be based upon a decision theory metric, whereby the reliability estimates of the various location cues are combined via a weighted average to identify the sound source. The model identifies gaps in the physiological knowledge base that apply to the methods involved in localization.

86. Searle, C. L., Braida, L. D., Cuddy, D. R., & Davis, M. F. (1975). Binaural pinna disparity: Another auditory localization cue. The Journal of the Acoustical Society of America, 57, 448-455.

This study measured pinna transformations in the left and right ear, for various frequencies and elevations in the median plane. The results showed significant differences between the transform functions for the two ears. These differences were assumed to be due to pinna asymmetry. The pinna transform disparities were evaluated as possible cues to aid auditory localization. It was concluded that these disparities are detectable and are used to aid localization in the median (and possibly the horizontal) plane.

87. Searle, C. L., Braida, L. D., Davis, M. F., & Colburn, H. S. (1976). Model for auditory localization. The Journal of the Acoustical Society of America, 60, 1164-1175.

A mathematical model of auditory localization was developed and validated against experimental findings. The model is based on the theory of signal detectability. Six cues for auditory localization were identified, including interaural time and amplitude differences (ITD and IAD, respectively), monaural head shadow, binaural pinna amplification response, monaural pinna amplification response, and shoulder bounce. ITD, IAD, and monaural head shadow provided cues in the horizontal plane, while the binaural pinna response, monaural pinna response, and shoulder bounce provided cues in both the horizontal and vertical plane. The binaural cues were assumed to be absolute cues to localization, in that no a priori information was required, while the monaural cues were assumed to be relative cues to location. The localization model, for both the horizontal and vertical plane, assumed that head movements were restrained, the cues were independent, scalar, gaussian random variables, and that the subject's response criterion was unbiased. The decision vector, reduced to a scalar quantity due to the above assumptions, determined source location based upon the weighted average of the six cues. The decision judgement is based upon the observer's estimate of the decision scalar. The model was validated by comparing the model's predictions of standard error of location against the average error magnitude of the empirical results. A Chi-square goodness of fit test revealed that the model adequately fit the test data. The model does not account for localization with head movements, which, if included, should improve the estimates of localization accuracy.

88. Shaw, E. A. G. (1974a). The external ear. In W. D. Keidel & W. D. Neff (Eds.), Handbook of Sensory Psychology, Vol. 5 (pp. 455-490). New York: Springer-Verlag.

This article reviewed the acoustical properties of the pinna, ear canal, head, and torso. Data from several sources were combined to yield characteristic information on the pinna's sound pressure changes and interaural time and phase differences. Anthropometric data on ear canal, pinna flange, and concha in human subjects, as well as other mammals' ear characteristics, were given. A review of physiological noise in the auditory system was also undertaken. In general, the paper gives a comprehensive synthesis on the effects of the external ear on incoming sound sources.

89. Shaw, E. A. G. (1974b). Tranformation of sound pressure level from the far field to the eardrum in the horizontal plane. The Journal of the Acoustical Society of America, 56, 1848-1861.

Twelve studies which measured pinna transformations were compared. The procedures employed to collapse data over studies were discussed, and families of curves were presented. Differences between studies and irregularities in the curves were identified.

90. Shaw, E. A. G. (1982a). External ear response and sound localization. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 30-41). Groton, CT: Amphora Press.

Research related to the role of the pinna in sound localization was reviewed. Specifically, the means of spectral encoding (i.e., filtering models), the role of pinna disparity cues, and intersubject differences were identified as unresolved problems. Comparisons were also made between human and KEMAR manikin eardrum response curves.

91. Shaw, E. A. G. (1982b). 1979 Rayleigh medal lecture: The elusive connection. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Applications (pp. 13-29). Groton, CT: Amphora Press.

The work on pinna-based transformations of auditory stimuli was reviewed, and the analysis was extended to include how those transformations are modified by the characteristics of the middle ear. It was found that the tympanic membrane and ossicles tend to reduce the energy available at the oval window. The inner ear receives only 10% of the energy available from the sound field.

92. Shaw, E. A. G., & Teranishi, R. (1968). Sound pressure generated in an external ear replica and real human ears by a nearby point source. The Journal of the Acoustical Society of America, 44, 240-349.

Sound pressure changes were measured at various azimuths using a rubber model of the pinna, concha, and ear canal. Measurements of sound pressure were made for both open canal and blocked ear canal conditions. The results were identical for both conditions up to 12 kHz. Five acoustic pressure gain maxima are shown which vary in frequency as a function of azimuth. Comparisons were also made using real ears, and were in good agreement. The paper provides comprehensive data on pinna effects on pressure over the entire range of audible frequencies.

93. Shelton, B. R., Rodger, J. C., & Searle, C. C. (1982). The relation between vision, head motion and accuracy of free-field auditory localization. The Journal of Auditory Research, 22, 1-7.

The role of vision and head motion in facilitating auditory localization was examined by restricting vision, head movement, or a combination of the two. Subjects' ability to accurately localize filtered noise, in the left front quadrant, was found to be subject to an interaction between head motion and vision: head motion only improves performance if normal vision is available. It was concluded that the major component in facilitating auditory localization is visually fixating the apparent locus of the auditory stimulus.

94. Shelton, B. R., & Searle, C. L. (1980). The influence of vision on the absolute identification of sound source position. Perception and Psychophysics, 28, 589-596.

Two experiments were designed to investigate the influence of vision on auditory localization. Experiment 1 varied speaker orientation (front, side, back, and vertical positions) and the availability of vision. The effects of vision improved performance in all conditions except vertical. Even when the speaker positions were out of sight (back), the availability of vision improved performance. It was suggested that vision, in this condition, provided a perceptual anchor or frame of reference. Experiment 2 investigated the role of several factors in visual facilitation of auditory localization. The results suggested that vision facilitates performance when visual and auditory information is correlated.

95. Sivian, L. J., & White, S. D. (1933). On minimum audible sound fields. The Journal of the Acoustical Society of America, 5, 288-321.

This study measured the minimum audible field (MAF), both monaurally and binaurally, over a frequency range of 100 to 15000 Hz, and over various azimuths. The MAF threshold measure is obtained relative to free field measurements, and indicates pressure thresholds at the observer's head. In contrast, the minimum audible pressure (MAP) is obtained in terms of pressure thresholds at the observer's eardrum. In general, MAF values relate to the usual mode of hearing, while MAP values relate to ear mechanics (e.g., eardrum displacement). A U-shaped function of MAF curves was found in both monaural and binaural conditions at 0 degrees azimuth. The pressure necessary for detection increased up to approximately 4000 Hz, and then decreased. Pressure decreased as azimuth increased and reached a maximum when the open ear faced the speaker. The pressure decreases tended to get larger as frequency increased. Comparisons of MAP and MAF curves identified lower thresholds, across the frequency range, for MAP values relative to MAF values. Several problems in obtaining MAP values were discussed, mainly concerned with insufficient 1930's technology.

96. Stewart, G. W. (1922). The intensity logarithmic law and the difference of phase effects in binaural audition. <u>Psychological Monographs</u>, <u>31</u>, 30-44.

This article evaluated the constant k, computed for each subject, in the intensity logarithmic law: $\Theta = k\text{Log}_e\ I_R/I_L$. The law relates the angular displacement of a sound source (Θ) to the intensity of the source at the right and left ears (I_R and I_L, respectively). The constant k was found to vary considerably between subjects, and within an individual, k decreased as frequency increased. The perceived angular displacement of a sound source was also measured as phase relationships between the two tones and the ear were varied. There tended to be an upper frequency limit of 1260 Hz through which the alteration of phase produced a fused auditory image. These effects have been well documented as localization cues, and now are referred to as interaural amplitude and time or phase difference cues.

97. Thurlow, W. R., Mangels, J. W., & Runge, D. S. (1967). Head movements during sound localization. <u>The Journal of the Acoustical Society of America</u>, <u>42</u>, 489-493.

The Thurlow and Runge (1967) study of induced head motion was extended in this paper through the evaluation of subject-induced head movement. Subjects were presented with noise signals, in either the horizontal or vertical plane, and their head movements used to localize the sound were filmed. It was predicted that rotational and pivot movements would produce the greatest change in interaural time and intensity cues, and would, therefore, be used most often. In terms of frequency, subjects tended to employ rotational, rotational-tip-and-pivot, and rotational-pivot movements most often. Rotational movements are larger (in degrees) than tip on pivot movements. Subjects also tended to rotate and tip their heads towards the sound source, but usually not enough to directly face the energized speaker. It was concluded that most, but not all, subjects will naturally move their heads to aid in localizing sound sources.

98. Thurlow, W. R., & Runge, P. S. (1967). Effect of induced head movements on localization of direction of sounds. <u>The Journal of the Acoustical Society of America</u>, <u>42</u>, 480-488.

Three experiments investigated the effects of induced head motion in the horizontal and vertical planes. Five types of movement were evaluated: no movement, rotation, pivot, tip, rotate-pivot and free head movement. Head movement was directly controlled by the experimenters. The results indicated a reduction in front-back reversals and in horizontal localization accuracy with movement. Vertical location accuracy improved only slightly with head movement. Rotational movement provided greater accuracy than pivot and tip movements. It was concluded that Wallach's (1940) theoretical discussion on the importance of head movements was empirically supported.

99. Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. Journal of Experimental Psychology, 27, 339-368.

The effects of head movement (either actual or perceived) on auditory localization was demonstrated. In one experiment the movement of a subject's head controlled the sound source such that it swung through an angle twice the size of the head movement. The perception of sound source location was approximately a mirror image (180° reversed) from the actual location. Only when the source remained constant did the subject accurately localize the source. Passive and apparent movement produced identical results.

100. Warren, D. H. (1970). Intermodality interactions in spatial localization. Cognitive Psychology, 1, 114-133.

A series of experiments investigated the role of vision and eye movements in auditory localization. In general, the results indicated that visual input improved auditory localization only when a structured visual environment was available. In the dark, or with eyes closed, auditory localization was poorer than with eyes open. The availability of a visual environment apparently allows the listener to map the auditory stimulus onto a visual representation of the environment.

101. Watkins, A. J. (1978). Psychoacoustical aspects of synthesized vertical locale cues. The Journal of the Acoustical Society of America, 63, 1152-1165.

Pinna transformations were simulated through a two variable latency delay and add process. Using Batteau's (1967, 1968) data as parameters, the computer simulation was able to produce apparent movement in the vertical plane similar to previous findings. A quantitative model of location decoding was proposed. The model is based upon spectral-pattern recognition (i.e., autocorrelation). The predictions of the model were consistant with the apparent elevation of the stimuli.

102. Watkins, A. J. (1982). The monaural perception of azimuth: A synthesis approach. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 194-206). Groton, CT: Amphora Press.

This article expanded upon Batteau's (1967) work on the directional band theory of sound localization. The theory proposes that the pinna produce two distinct echoes which vary in latency and are used as cues to define the sound source location. Three experiments were performed, using monaural cues presented through a single earphone. The results showed that subjects perceive changes in azimuth dependent upon the interaction of the two time parameters associated with the two echoes. The results supported the predictions made by directional band theory.

103. Weiner, F. M. (1947). On the diffraction of a progressive sound wave by the human head. The Journal of the Acoustical Society of America, 19, 143-146.

An earlier paper by Wiener and Ross (1946) investigated changes in sound pressure from the entrance to the ear canal to the eardrum. In general, sound pressure increased, especially at higher frequencies, at the eardrum relative to the ear canal. This data was then combined with data from free field and ear canal measurements, to obtain a measure of pressure changes due to both head diffraction and ear canal factors at the eardrum. Measurements were taken over various frequencies and azimuthal positions. It was concluded that pressure changes at the eardrum are sufficient, especially at higher frequencies, to encode azimuthal changes.

104. Weinrich, L. (1982). The problem of front-back localization in binaural hearing. In O. J. Pedersen and T. Poulsen (Eds.), Binaural Effects in Normal and Impaired Hearing (Scandinavian Audiological Supplement, No. 15, pp. 135-145).

Intensity measurements of pinna transfer functions were made. It was concluded that changes in the functions were available as a cue, without head movements, to reduce directional inversions. Using these transfer functions, it was found that subjects were able, whether or not the signal characteristics were modified, to localize the sound source without front-back inversions. These modifications, in the 500 - 7000 Hz range, consisted of manipulating the tonal characteristics of the noise bursts such that the transfer functions were taken from one orientation (i.e., speaker location) and presented at another location.

105. Welsh, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. Psychological Bulletin, 88, 638-667.

A review of the intersensory bias literature was performed with regards to sensory modality, the type of discrepancy (either perceptual or cognitive), and response variables. The previous theories are unable to account for all the results of intersensory bias and perceptual adaptation. A model is proposed to account for the effects of stimulus situations, modality characteristics, cognitive factors, and response outcomes. The model is broad enough to cover the known occurrences of intersensory bias.

106. Wilbanks, W. A. (1983). Masking of the signal for lateralization of tones. Bulletin of the Psychonomic Society, 21, 138-140.

This study evaluated the change in interaural phase necessary to detect lateralization of tones under conditions of different source frequency and interaural correlations of background wide band noise. Using Zwislocki and Feldman's (1956) data as a baseline for no noise (≈2° phase difference), the results showed that as the tonal frequency increased phase adjustments increased, and as the interaural noise correlation decreased phase adjustments increased. With noise correlation equaling +1.0 for a 500 Hz tone, the phase adjustment increased from Zwislocki and Feldman's no noise condition to 8°. It is concluded that both the frequency and interaural noise correlations are important variables in attempting to lateralize a tonal signal imbedded in background noise.

107. Wright, D., Hebrank, J. H. & Wilson, B. (1974). Pinna reflections as cues for localization. The Journal of the Acoustical Society of America, 56, 957-962.

This study investigated the monaural difference threshold of time delays. Batteau (1967) proposed that pinna reflections introduced time delays of 10 to 100 microseconds, which were dependent upon sound source direction. It was found that delay times as little as 20 microseconds were recognizable when the amplitude ratio of the two signals were greater than .67, and that JND values were consistent with previously obtained minimum audible angles for monaural localization. These findings support Batteau's theory.

108. Zwislocki, J., & Feldman, R. S. (1956). Just noticeable differences in dichotic phase. The Journal of the Acoustical Society of America, 28, 860-864.

Phase differences, in a dichotic listening task, were directly manipulated to assess the relationship between changes in phase and the more commonly obtained changes in times. Just noticeable differences (JNDs) were obtained from three subjects, using a forced choice paired comparison method in which the intracranial location of a standard was judged relative to a test stimulus. Both frequency and intensity were varied. The results indicated that phase discrimination between the two stimuli was poorest at both high and low intensities, and that as frequency increased the JND for phase difference increased. Interaural phase differences on the order of 2° were perceived as a shift in the location of the test stimulus. These results were consistent with other findings, and supported the notion that phase differences in localization are effective only at the lower frequencies.

TOPICAL INDEX

The numbers under each topic refer to the annotation number in the appendix.

8.0 REFERENCES

Alekseenko, N. Y. (1983). Role of movements of different types in spatial hearing activity. Human Physiology, 8, 240-243.

Andreassi, J. L., & Greco, J. R. (1975). Effects of bisensory stimulation on reaction time and the evoked cortical potential. Physiological Psychology, 3, 189-194.

Batteau, D. W. (1967). The role of the pinna in human localization. Proceedings of the Royal Society (pp. 158-180). London, England.

Batteau, D. W. (1968). Listening with the naked ear. In S. J. Freedman (Ed.), The Neuropsychology of Spatially Oriented Behavior (pp. 109-134). Homewood, IL: Dorsey Press.

Batteau, D. W., Plante, R. L., Spencer, R. H., & Lyle, W. E. (1963). Localization of sound: Part 3. A new theory of human audition (TP3109, part 3). China Lake, CA: U. S. Naval Ordinance Test Station.

Batteau, D. W., Plante, R. L., Spencer, R. H., & Lyle, W. E. (1965). Localization of sound: Part 5. Auditory perception (TP3109, part 5). China Lake, CA: U. S. Naval Ordinance Test Station.

Bauer, B. B. (1961). Phasor analysis of some stereophonic phenomena. The Journal of the Acoustical Society of America, 33, 1536-1539.

Bekesy, G. von (1949). The moon illusion and similiar auditory phenomena. American Journal of Psychology, 62, 540-552.

Bekesy, G. von (1960). Experiments in Hearing. New York: Wiley.

Biguer, B., Jeannerod, M., & Prablanc, C. (1982). The coordination of eye, head, and arm movements during reaching at a single visual target. Experimental Brain Research, 46, 301-304.

Blauert, J. (1969/1970). Sound localization in the median plane. Acustica, 22, 205-213.

Blauert, J. (1971). Localization and the law of the first wavefront in the median plane. The Journal of the Acoustical Society of America, 50, 466-470.

Blauert, J. (1981). Lateralization of jittered tones. The Journal of the Acoustical Society of America, 70, 694-698.

Blauert, J. (1982). Binaural localization. In O. J. Pedersen and T. Poulsen (Eds.), Binaural Effects in Normal and Impaired Hearing (Scandinavian Audiological Supplement No. 15, pp. 7-20).

Blauert, J. (1983). Spatial Hearing: The Psychophysics of Human Sound Localization. Cambridge, MA: The MIT Press.

Briggs, R., & Perrott, D. R. (1972). Auditory apparent movement under dichotic listening conditions. Journal of Experimental Psychology, 92, 83-91.

Burkhard, M. D., & Sachs, R. M. (1975). Anthopometric manikin for acoustic research. The Journal of the Acoustical Society of America, 58, 214-222.

Burtt, H. E. (1917). Auditory illusions of movement: A preliminary study. Journal of Experimental Psychology, 2, 63-75.

Butler, R. A. (1969a). Monaural and binaural localization of noise bursts vertically in the median sagittal plane. The Journal of Auditory Research, 3, 230-235.

Butler, R. A. (1969b). On the relative usefulness of monaural and binaural cues in locating sound in space. Psychonomic Science, 17, 245-246.

Butler, R. A. (1970). The effect of hearing impairment on locating sound in the vertical plane. International Audiology, 9, 117-126.

Butler, R. A., & Belendiuck, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. The Journal of the Acoustical Society of America, 61, 1264-1269.

Butler, R. A., & Flannery, R. (1980). The spatial attributes of stimulus frequency and their role in monaural localization of sound in the horizontal plane. Perception and Psychophysics, 28, 449-457.

Butler, R. A., & Helwig, B. E. (1983). The spatial attributes of stimulus frequency in the median sagittal plane and their role in sound localization. American Journal of Otolaryngology, 4, 165-173.

Canevet, G., Germain, R., & Scharf, B. (1979). Effect of one tone burst on the localization of a second tone burst. The Journal of the Acoustical Society of America, 65, S121.

Canevet, G., Germain, R., & Scharf, B. (1980). Sound localization in the presence of a masking sound. Acustica, 46, 96-99.

Chason, L. R., McFarland, T. P., & Aldrich, T. B. (1971). Auditory Effects on Spatial Orientation (Report No. 71-10). Colorado Springs, CO: United States Air Force Academy. (AD 737351).

Coleman, P. D. (1962). Failure to localize the source distance of an unfamiliar sound. The Journal of the Acoustical Society of America, 34, 345-346.

Coleman, P. D. (1963). An analysis of cues to auditory depth perception in free space. Psychological Bulletin, 60, 302-315.

Coleman, P. D. (1968). Dual role of frequency spectrum in determination of auditory distance. The Journal of the Acoustical Society of America, 44, 631-632.

Dirks, D. D., & Gilman, S. (1979). Exploring azimuth effects with an anthropometric manikin. The Journal of the Acoustical Society of America, 66, 696-701.

Doll, T. J., Folds, D. J., & Leiker, L. A. (1984). Auditory Information Systems In Military Aircraft: Current Configurations Versus the State of the Art (Report No. USAFSAM-TR-84-15). Atlanta, GA: Georgia Institute of Technology.

Doll, T. J., Gerth, J. M., Engelman, W. R., & Folds, D. J. (1985). Applications of Simulated Auditory Localization in Head-Coupled Control/Display Systems. Interim report to the Armstrong Aerospace Medical Research Laboratory. Georgia Institute of Technology, Atlanta, GA.

Durlach, N. I. (1964). Note on binaural masking-level differences as a function of the interaural correlation of the masking noise. The Journal of the Acoustical Society of America, 36, 1613-1617.

Durlach, N. I. (1972). Binaural signal detection: Equalization and cancellation theory. In J. V. Tobias (Ed.), Foundations of Modern Auditory Theory, Vol. II (pp. 371-462). New York: Academic Press.

Durlach, N. I., & Colburn, H. S. (1978). Binaural phenomenon. In E. C. Carterette and M. P. Friedman (Eds.), Handbook of Perception: Vol. IV (pp. 365-466). New York: Academic Press.

Easton, R. D. (1983). The effect of head movements on visual and auditory dominance. Perception, 12, 63-70.

Feddersen, W. E., Sandel, T. T., Teas, D. L., & Jeffress, L. A. (1957). Localization of high frequency tones. The Journal of the Acoustical Society of America, 29, 988-991.

Firestone, F. A. (1930). The phase difference and amplitude ratio at the ears due to a source of pure tones. The Journal of the Acoustical Society of America, 2, 260-270.

Flannery, R., & Butler, R. A. (1981). Spectral cues provided by the pinna for monaural localization in the horizontal plane. Perception and Psychophysics, 29, 438-444.

Forbes, T. W. (1946). Auditory signals for instrument flying. Journal of the Aeronautical Sciences, May, 255-258.

Freedman, S. J., & Fisher, H. G. (1968). The role of the pinna in auditory localization. In S. J. Freedman (Ed.), The Neuropsychology of Spatially Oriented Behavior (pp. 135-152). Homewood, IL: Dorsey Press.

Freedman, S. J., & Pfaff, D. W. (1962). The effect of dichotic noise on auditory localization. The Journal of Auditory Research, 2, 305-310.

Freides, D. (1974). Human information processing and sensory modality: Cross-modal functions, information, complexity, memory, and deficit. Psychological Bulletin, 81, 284-310.

Gardner, M. B. (1969a). Distance estimation of 0° or apparent 0°-oriented speech signals in anechoic space. The Journal of the Acoustical Society of America, 45, 47-53.

Gardner, M. B. (1969b). Image fusion, broadening, and displacement in sound localization. The Journal of the Acoustical Society of America, 46, 339-349.

Gardner, M. B. (1973). Some monaural and binaural facets of median plane localization. The Journal of the Acoustical Society of America, 54, 1489-1495.

Gardner, M. B., & Gardner, R. S. (1973). Problem of localization in the median plane: Effect of pinnae cavity occlusion. The Journal of the Acoustical Society of America, 53, 400-408.

Garner, W. R. (1949). Auditory signals. In A survey report on human factors in undersea warfare (pp. 201-217). Washington, DC: National Research Council.

Gaskell, H., & Henning, G. B. (1979). The effect of noise on time/intensity trading in lateralization. The Journal of the Acoustical Society of America, 65, S121-S122.

Gatehouse, R. W. (1982a). Introduction. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 4-12). Groton, CT: Amphora Press.

Gatehouse, R. W. (1982b). Summary: New directions. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 267-270). Groton, CT: Amphora Press.

Geldard, F. A. (1984). Is there a lesson for audition in tactual localization. Advances in Audiology, 1, 117-127.

Gopher, D. (1973). Eye movement patterns in selective listening tasks of focused attention. Perception and Psychophysics, 14, 259-264.

Gotoh, T. (1982). Can the acoustic head related transfer function explain every phenomeon in sound localization. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 244-248). Groton, CT: Amphora Press.

Grantham, D. W., & Wightman, F. L. (1978). Detectability of varying interaural temporal differences. The Journal of the Acoustical Society of America, 63, 511-523.

Greene, D. C. (1968). Comments on "Perception of the range of a sound source of unknown strength". The Journal of the Acoustical Society of America, 44, 634.

Green, D. M. (1971). Temporal auditory acuity. Psychological Review, 78, 540-551.

Harris, J. D. (1972). A florilegium of experiments on directional hearing. Acta Oto-Laryngologica, Supplement 298.

Harris, J. D., & Sergeant, R. L. (1971). Monaural/binaural minimum audible angles for a moving sound source. Journal of Speech Hearing Research, 14, 618-629.

Hebrank, J., & Wright D. (1974a). Are two ears necessary for localization of sound sources on the median plane. The Journal of the Acoustical Society of America, 56, 935-938.

Hebrank, J., & Wright D. (1974b). Spectral cues used in the localization of sound sources on the median plane. The Journal of the Acoustical Society of America, 56, 1829-1834.

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. Journal of Experimental Psychology, 63, 289-293.

Hirsh, H. R. (1968). Perception of the range of a sound source of unknown strength. The Journal of the Acoustical Society of America, 43, 373-374.

Holt, R. E., & Thurlow, W. R. (1969). Subject orientation and judgement of distance of a sound source. The Journal of the Acoustical Society of America, 46, 1584-1585.

Ito, Y., Thompson, C. L., & Colburn, H. S. (1979). Interaural time discrimination in noise. The Journal of the Acoustical Society of America, 65, 5121.

Jeffress, L. A. (1972). Binaural signal detection: Vector theory. In J. V. Tobias (Ed.), Foundations of Modern Auditory Theory, Vol. II (pp. 351-368). New York: Academic Press.

Jones, B. (1975). Visual facilitation of auditory localization in school children: A signal detection analysis. Perception and Psychophysics, 17, 217-220.

Jones, B., & Kabanoff, B. (1975). Eye movements in auditory space perception. Perception & Psychophysics, 17, 241-245.

King, W. G., & Laird, D. A. (1930). The effect of noise intensity and pattern on locating sound. The Journal of the Acoustical Society of America, 2, 99-102.

Kock, W. E. (1950). Binaural localization and masking. The Journal of the Acoustical Society of America, 22, 801-804.

Kuhn, G. F. (1977). Model for the interaural time differences in the azimuthal plane. The Journal of the Acoustical Society of America, 62, 157-167.

Kuhn, G. F. (1982). Towards a model for sound localization. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 51-64). Groton, CT: Amphora Press

Lambert, R. M. (1974). Dynamic theory of sound-source localization. The Journal of the Acoustical Society of America, 56, 165-171.

Levy, E. T., & Butler, R. A. (1978). Stimulus factors which influence the perceived externalization of sound presented through headphones. The Journal of Auditory Research, 18, 41-50.

Lounsbury, B. F., & Butler, R. A. (1979). Estimation of distances of recorded sounds presented through headphones. Scandinavian Audiology, 8, 145-149.

Loveless, N. E., Brebner, J., & Hamilton, P. (1970). Bisensory presentation of information. Psychological Bulletin, 73, 161-199.

Mastroianni, G. R. (1982). The influence of eye movements and illumination of auditory localization. Perception & Psychophysics, 31, 581-584.

Mathiesen, A. (1931). Apparent movement in auditory perception. Psychological Monographs, 41, 74-131.

Matin, L. (1982a). Visual and auditory localization: Normal and abnormal relations. In G. T. Chisum & P. E. Morway (Eds.), Research Program Review: Aircrew Physiology (Report No. NADC-82232-60). Warminster, PA: Naval Air Development Center.

Matin, L. (1982b). Visual localization and eye movements. In A. H. Wertheim, W. A. Wagenaar, & H. W. Leibowitz (Eds.), Tutorials on Motion Perception (pp. 101-156). New York: Plenum Press.

Matin, L., Picoult, E., Stevens, J. K., Edwards, M. W., Jr., Young, D., & MacArthur, R. (1982). Oculaparalytic illusion: Visual-field dependent spatial mislocalization by humans partially paralyzed with curare. Science, 216, 198-201.

Matin, L., Stevens, J. K., & Picoult, E. (1983). Perceptual consequences of experimental extraocular muscle paralysis. In A. Heim and M. Jeannerod (Eds.), Spatially Oriented Behavior (pp. 243-262) New York: Springer-Verlag.

McFadden, D. (1969). Lateralization and detection of a tonal signal in noise. The Journal of the Acoustical Society of America, 45, 1505-1509.

McFadden, D., & Pasanen, E. G. (1975). Binaural beats at high frequencies. Science, 190, 394-396.

McFadden, D., & Pasanen, E. G. (1976). Lateralization at high frequencies based on interaural time differences. The Journal of the Acoustical Society of America, 59, 634-639.

McNulty, J. A. (1982). Current aspects of underwater localization. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 250-266). Groton, CT: Amphora Press.

Mehrgardt, S., & Mellert, V. (1977). Transformation characteristics of the external human ear. The Journal of the Acoustical Society of America, 61, 1567-1576.

Metcalfe, J., Glavanov, D., & Murdock, M. (1981). Spatial and temporal processing in the auditory and visual modalities. Memory & Cognition 9, 351-359.

Mills, A. W. (1958). On the minimum audible angle. The Journal of the Acoustical Society of America, 30, 237-246.

Mills, A. W. (1963). Auditory perception of spatial relations. In Proceedings of the International Congress on Technology and Blindness, Vol. 2 (pp. 111-139). New York: American Foundation for the Blind.

Mills, A. W. (1972). Auditory localization. In J. V. Tobias (Ed.), Foundations of Modern Auditory Theory, Vol. II (pp. 303-348). New York: Academic Press.

Molino, J. (1973). Perceiving the range of a sound source when the direction is known. The Journal of the Acoustical Society of America, 53, 1301-1304.

Morimoto, M., & Ando, Y. (1982). On the simulation of sound localization. In R. W. Gatehouse (Ed.), Localisation of Sound: Theory and Application (pp. 85-98). Groton, CT: Amphora Press.

Moushegian, G., & Jeffress, L. A. (1959). Role of inter-aural time and intensity differences in the lateralization of low-frequency tones (Report No. DRL-A-144) Austin, TX: The University of Texas. (AD-A031-955)

Mudd, S. A., & McCormick, E. J. (1960). The use of auditory cues in visual search task. Journal of Applied Psychology, 44, 184-188.

Mudd, S. A. (1965). Experimental evaluation of binary puretone auditory displays. Journal of Applied Psychology, 49, 112-121.

Mulligan, R. M., & Shaw, M. L. (1981). Attending to simple auditory and visual signals. Perception & Psychophysics, 30, 447-454.

Murch, G. M. (1973). Visual and Auditory Perception. New York: Bobbs-Merrill.

Musicant, A. D., & Butler, R. A. (1984a). The influence of pinnae-based spectral cues on sound localization. The Journal of the Acoustical Society of America, 75, 1195-1200.

Musicant, A. D., & Butler, R. A. (1984b). The psychophysical basis of monaural localization. Hearing Research, 14, 185-190.

Navon, D., & Gopher, D. (1979). On the economy of the human processing system. Psychological Review, 86, 214-255.

Nikerson, R. S. (1973). Intersensory facilitation of reaction time: Energy summation or preparation enhancement? Psychological Review, 80, 489-509.

Nordlund, B. (1962). Physical factors in angular localization. Acta Otolaryngology, 54, 75-93.

Norman, D., & Bobrow, D. (1975). On the data-limited and resource-limited processes. Cognitive Psychology, 7, 44-64.

Oatman, L. C. (1975). Simulataneous processing of bisensory information. Aberdeen, MD: U.S. Army Human Engineering Laboratory. (AD-A012 149)

Perrott, D. R. (1982). Studies in the perception of auditory motion. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 169-193) Groton, CT: Amphora Press.

Perrott, D. R., & Baars, B. (1974). Detection of interaural onset and offset disparities. The Journal of the Acoustical Society of America, 55, 1290-1292.

Perrott, D. R., & Elfner, L. F. (1968). Monaural localization. The Journal of Auditory Research, 8, 185-193.

Perrott, D. R., Mason, G., & Forbes, J. (1973). The instability of auditory perceptual experience in auditory autokinesis. The Journal of Auditory Research, 13, 80-86.

Perrott, D. R., & Musicant, A. D. (1977). Rotating tones and binaural beats. The Journal of the Acoustical Society of America, 61, 1288-1292.

Perrott, D. R., & Musicant, A. D. (1981). Dynamic minimum audible angle: Binaural spatial acuity with moving sound sources. Journal of Auditory Research, 21, 287-295.

Pick, H. L., Warren, D. H., & Hay, J. C. (1969). Sensory conflict judgments of spatial direction. Perception & Psychophysics, 6, 203-205.

Platt, B., & Warren, D. (1972). Auditory localization: The importance of eye movements and a textured visual environment. Perception & Psychophysics, 12, 245-248.

Pollack, I., & Rose, M. (1967). Effect of head movement on the localization of sounds in the equatorial plane. Perception and Psychophysics, 2, 591-596.

Posner, M. E. (1978) Chronometric Explorations of Mind. New Jersey: John Wiley & Sons.

Robinson, D. E., & Egan, J. P. (1974). Lateralization of an auditory signal in correlated noise and in uncorrelated noise as a function of signal frequency. Perception & Psychophysics, 15, 281-284.

Rodgers, C. A. (1981). Multidimensional localization: An investigation of possible pinnae cues. Dissertation Abstracts International, 42, 1811B.

Roffler, S., & Butler, R. A. (1968). Localization of tonal stimuli in the vertical plane. The Journal of the Acoustical Society of America, 43, 1260-1266.

Russell, G. (1976). The role of the pinna in monaural horizontal plane localization. Journal of Auditory Research, 16, 68-70.

Sandel, T. T., Teas, D. C., Feddersen, W. E., & Jeffress, L. A. (1955). Localization of sound from single and paired sources. The Journal of the Acoustical Society of America, 27, 842-852.

Scharf, B., Canevet, G., Buus, S., & Marchioni, A. (1982, May). Localization of noise by hearing impaired listeners. Paper presented at The 16th International Congress of Audiology, Helsinki, Finland.

Searle, C. L. (1982). A model of auditory localization: Peripheral constraints. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 42-50). Groton, CT: Amphora Press.

Searle, C. L., Braida, L. D., Cuddy, D. R., & Davis, M. F. (1975). Binaural pinna disparity: Another auditory localization cue. The Journal of the Acoustical Society of America, 57, 448-455.

Searle, C. L., Braida, L. D., Davis, M. F., & Colburn, H. S. (1976). Model for auditory localization. The Journal of the Acoustical Society of America, 60, 1164-1175.

Shaw, E. A. G. (1974a). The external ear. In W. D. Keidel & W. D. Neff (Eds.), Handbook of Sensory Psychology, Vol. 5 (pp. 455-490). New York: Springer-Verlag.

Shaw, E. A. G. (1974b). Transformation of sound pressure level from the far field to the eardrum in the horizontal plane. The Journal of the Acoustical Society of America, 56, 1848-1861.

Shaw, E. A. G. (1982a). External ear response and sound localization. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 30-41). Groton, CT: Amphora Press.

Shaw, E. A. G. (1982b). Rayleigh medal lecture: The elusive connection. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 13-29). Groton, CT: Amphora Press.

Shaw, E. A. G., & Teranishi, R. (1968). Sound pressure generated in an external-ear replica and real human ears by a nearby point source. The Journal of the Acoustical Society of America, 44, 240-249.

Shelton, B. R., Rodgers, J. C., & Searle, C. L. (1982). The relation between vision head motion and accuracy of free-field auditory localization. Journal of Auditory Research, 22, 1-7.

Shelton, B. R., & Searle, C. L. (1980). The influence of vision on the absolute identification of sound-source position. Perception & Psychophysics, 28, 589-596.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

Simpson, C. A. (1982, April). Speech I/O in the cockpit: Plane to pilot...pilot to plane. Design News, pp. 159-168.

Sivian, L. J. & White, S. D. (1933). On minimum audible sound fields. The Journal of the Acoustical Society of America, 5, 288-321.

Stevens, S. S., & Newman, E. B. (1936). The localization of actual sound sources. American Journal of Psychology, 48, 297-306.

Stewart, G. W. (1922). The intensity logarithmic law and the difference of phase effect in binaural audition. Psychological Monographs, 31, 30-44.

Tasaki, I. (1954). Nerve impulses in individual auditory nerve fibers of guinea pig. Journal of Neurophysiology, 17, 97-122.

Thurlow, W. R., Mangels, J. W., & Runge, P. S. (1967). Head movements during sound localization. The Journal of the Acoustical Society of America, 42, 489-493.

Thurlow, W. R., & Runge, P. S. (1967). Effect of induced head movements on localization of direction of sounds. The Journal of the Acoustical Society of America, 42, 480-488.

Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. Journal of Experimental Psychology, 27, 339-368.

Warren, D. H. (1970). Intermodality interactions in spatial localization. Cognitive Psychology, 1, 114-133.

Watkins, A. J. (1978). Psychoacoustical aspects of synthesized vertical locale cues. The Journal of the Acoustical Society of America, 63, 1152-1165.

Watkins, A. J. (1982). The monaural perception of azimuth: A synthesis approach. In R. W. Gatehouse (Ed.), Localization of Sound: Theory and Application (pp. 194-206). Groton, CT: Amphora Press.

Weiner, F. M. (1947). On the diffraction of a progressive sound wave by the human head. The Journal of the Acoustical Society of America, 19, 143-146.

Weinrich, L. (1982). The problem of front-back localization in binaural hearing. In O. J. Pedersen and T. Poulsen (Eds.), Binaural Effects in Normal and Impaired Hearing (Scandinavian Audiological Supplement No. 15, pp. 135-145).

Welsh, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. Psychological Bulletin, 88, 638-667.

Werkowitz, E. (1981, May). Ergonimic considerations for the cockpit applications of speech generation technology. Proceedings of the Voice-Interactive Systems: Applications and Payoffs Symposium. Naval Air Development Center.

Wettschureck, R. G. (1973). The absolute difference limen of directional perception in the median plane under conditions of both natural hearing and hearing with artificial-head-system. Acustica, 28, 197-208.

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), Attention and performance (Vol. VIII). Englewood Cliffs, NJ: Erlbaum.

Wickens, C. D., Sandry, D. L., & Vidulich, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output. Human Factors, 25, 227-248.

Wilbanks, W. A. (1983). Masking of the signal for lateralization of tones. Bulletin of the Psychonomic Society, 21, 138-140.

Wright, D., Hebrank, J. H., & Wilson, B. (1974). Pinna reflections as cues for localization. The Journal of the Acoustical Society of America, 56, 957-962.

Yost, W. A., Wightman, F. L., & Green, D. M. (1971). Lateralization of filtered clicks. The Journal of the Acoustical Society of America, 50, 1526-1531.

Zwicker, E. (1973). Temporal effects of psychoacoustic excitation. In A. R. Moller (Ed.), Basic Mechanisms in Hearing (pp. 809-824). London: Academic Press.

Zwislocki, J. J., & Feldman, R. S. (1956). Just noticeable differences in dichotic phase. The Journal of the Acoustical Society of America, 28, 860-864.

U S GOVERNMENT PRINTING OFFICE 1988 - 748-001/40221